Journal of Korean
Data & Information Science Society
2007, Vol. 18, No. 3, pp. 773 ~ 782

A Study on the Monitoring of Reject Rate in High Yield Process

Nam, Ho Soo[1)]

Abstract

The statistical process control charts are very extensively used for monitoring of process mean, deviation, defect rate or reject rate. In this paper we consider a control chart to monitor the process reject rate in the high yield process, which is based on the observed cumulative probability of the number of items inspected until r defective items are observed. We first propose selection of the optimal value of r in the CPC-r charts, and also consider the usefulness of the chart in high yield process such as semiconductor or TFT-LCD manufacturing process.

Keywords : CCC-r Chart, CPC-r Chart, High Yield Process

1. Introduction

The real time process monitoring in statistical process control is very important for prevention of mass disaster such as large amount of defective. Recently due to rapid growth of computer aided manufacturing system and process monitoring systems, the reject rate of the process, so called, fraction defectives of the process become very low. Especially centering around electronic industries, the six sigma management and zero defect thinking of production make the classical control charts such as p- or np- chart no more efficient and useful for monitoring of the process reject rate.

Usually p-chart used for monitoring of the reject rate of the process is based on the binomial distribution. In conventional Shewhart-type p-chart control limits (UCL and LCL) are given as follows:

$$\bar{p} \pm 3\sqrt{\bar{p}(1-\bar{p})/n_i}\,,$$

1) Associate Professor, Department of System & Management Engineering, Dongseo University, Busan, 617-716, Korea
E-mail : hsnam@dongseo.ac.kr

where n_i is size of ith subgroup and $\bar{p}$ is mean reject rate of the process. These control limits by the 3 times estimator of standard error of plotting statistics are based on approximation of binomial distribution to normal distribution. However the normal approximation of binomial is not accurate at the tail parts of distribution (Chan et al. 2002), and it requires tremendous subgroup size for nice normal approximation when the process reject rate is low.

For example, if average reject rate is about 100PPM(parts per million) then a meaningful subgroup size is at least above 50000, since the distribution of the number of rejected items is skewed to right.

By these reasons, Calvin(1983), Xie and Goh(1992, 1997) discussed an alternative process monitoring method based on the cumulative counts of conforming called CCC chart, which is derived from geometric distribution. The CCC-chart is very useful monitoring tool of low defect manufacturing processes. In succession the CCC-r chart was discussed by many other researchers.

The CCC chart is a special case of CCC-r(r=1) chart. Usually CCC-r chart is more powerful than CCC chart, since the CCC-r chart with large r requires large sample size. But the CCC-r chart takes long time for the monitoring system to give a warning. Therefore the decision of r in CCC-r chart is an important problem for the construction of efficient process monitoring system.

2. Overview of the CCC-r and CPC-r Charts for High Yield Process

Let X_i be an independently and identically distributed geometric random variable with common reject rate p. Then the sum of these variables N is defined as follows:

$$N = \sum_{i=1}^{r} X_i \sim Nb(r,p)$$

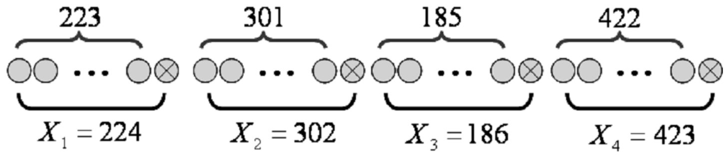
That is, the random variable N is distributed as negative binomial with parameter r and p, and N means the cumulative number of inspected until observing $r(\geq 1)$ non-conforming ones. The probability function of N is also given by

$$P(N=n) = \binom{n-1}{r-1} p^r (1-p)^{n-r},\ n = r, r+1, \cdots$$

where p represents the fraction defectives or reject rate of the process being monitored. The expectation, variance and standard error of the random variable N is obtained as follows.

$$E(N) = \frac{r}{p}; \quad Var(N) = \frac{r(1-p)}{p^2}; \quad SE(N) = \sqrt{\frac{r(1-p)}{p^2}}$$

This situation is described as figure 2.1. The figure shows an illustrative example of geometric and negative binomial distribution.



[Figure 2.1] Model of negative binomial distribution

In figure 2.1, the icon ○ means conforming item and the icon ⊗ means non-conforming(rejected) item. In this example X_i , $i=1,2,3,4$ distributed as geometric and the sum of X_i, say N, means negative binomial distribution with $r=4$.

In the figure, $X_1=224, X_2=302, X_3=186, X_4=423$ and $N=X_1+X_2+X_3+X_4=1135$ that is, the number of inspected items until 4 non-conforming items is 1135. In this case we can estimate the fraction defective p as

$$\hat{p}=\frac{r}{N}=\frac{4}{1135}=0.00352$$

In the CCC-r chart, the plotting points are N_j, $j=1,2,3,\cdots$, and the control lines based on the probability limits as defined as follows (Xie, et al.(1998)):

$$F(UCL_r,r,p)=\sum_{n=r}^{UCL_r}P(N=n)=1-\frac{\alpha}{2}$$

$$F(CL_r,r,p)=\sum_{n=r}^{CL_r}P(N=n)=0.5$$

$$F(LCL_r,r,p)=\sum_{n=r}^{LCL_r}P(N=n)=\frac{\alpha}{2}$$

where UCL_r, CL_r and LCL_r are upper control limit, center line and lower control limit respectively in CCC-r chart. The function means the cumulative distribution function, and α means the false alarm probability level when the process is in control.

While the cumulative probability control(CPC) chart is discussed by Chan, et.al.(2002). The CPC chart is a control chart in which the observed cumulative probability of random variate is plotted against the time series or sample number. In the CPC chart, the LCL, CL and UCL at false alarm rate α are always as follows:

$$LCL = \frac{\alpha}{2}$$
$$CL = 0.5$$
$$UCL = 1 - \frac{\alpha}{2}$$

In this paper, we propose the CPC-r chart, which is a generalized version of CPC chart such as generalization of the CCC chart to CCC-r chart. That is, the CPC-r chart uses the cumulative probability of the count instead of the cumulative count in CCC-r chart.

If $N_j\ (j = 1,2,3,\cdots)$ is the number of items inspected until r non-conforming items are observed, the cumulative probability is defined by as follows:

$$U_j = F(N_j) = \sum_{n=r}^{N_j} P(N = n)$$

The random variable U_j follows uniform distribution in range (0, 1), and that is the plotting point in CPC-r chart. In CPC-r chart, large value of U_j means large value of N_j, and it means the improvement of the process. But for small value of U_j, we know the process is deteriorated.

3. A selection of r in CPC-r chart

Usually the fraction defective p is estimated by $\frac{r}{N}$. That is, the process fraction defective is estimated by sample fraction defective. In high yield process the value of N is not small. For example, if the fraction defective p is about 0.1%, 1000PPM then we need at least a sample size of near 5000. In this case the classical p-chart do not perform its original role, real time monitoring of the process characteristics.

The control chart for monitoring of process fraction defective p is regarded as testing the following hypothesis.

$$H_0 : p = p_0 \quad v.s. \quad H_1 : p \neq p_0$$

Usually we conclude the test as more powerful one having high power and having same level for fixed sample size. However, if the sample sizes of two tests are not same, how can we compare which test is more powerful? If a test attains uniformly high power against various alternatives, and the test is based on more large sample size than the other test, can we still the test is more powerful than the other test? It is natural that for fixed significant level, the more sample size increases, the more power of test also increases.

We have a dilemma for the choice of optimal value of r in CCC-r chart, since the power of detection of CCC-r chart with large value of r always gets higher than other CCC-r chart with small value of r. However, the CCC-r chart with large r requires tremendous sample size, and it takes long times of interval to give a warning to monitoring system in high yield process.

The finite sample relative efficiency of two tests T_{N_1}, T_{N_2}, which is a useful measure of the performance of the tests, is defined by

$$eff(T_{N_1}, T_{N_2}) = \frac{N_2}{N_1}$$

where N_1, N_2 are sample size of the tests T_{N_1}, T_{N_2} respectively that the two tests achieve a power of at least β against an alternative for a fixed α-level.

Now we define a new measure of relative performance of two tests, which is based on the power of detection for alternatives and the relative efficiency of tests. When the sampling inspection costs per plotting a point equal in each test T_{N_1}, T_{N_2}, the performance of two tests is defined as follows:

$$perf(T_{N_1}, T_{N_2}) = \frac{N_2}{N_1} \cdot \frac{\beta(p_1, N_1)}{\beta(p_1, N_2)}$$

where $\beta(p_1, N_i)$ is power of detection of the test T_{N_i} when $p = p_1$, i.e., fraction defective of process is changed p_0 to p_1, and it also satisfies the significant level of tests $\beta(p_0, N_i) \le \alpha \quad (i = 1, 2)$.

If there are k tests to compare, we can modify the performance of tests as follows:

$$perf(T_{N_i}) = \frac{\max N_j}{N_i} \cdot \beta(p_1, N_i), \qquad i = 1, 2, \cdots, k,$$

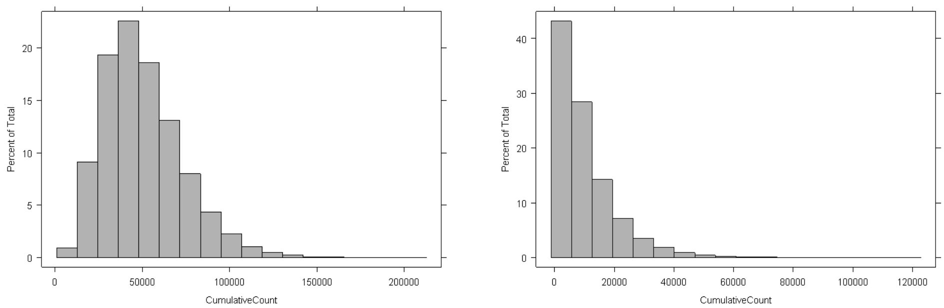
where N_i is sample size of the test T_{N_i}. That is, the performance of the test is proportional to the power of detection and inversely proportional to the sample size. Therefore for the test having maximum value of performance for various alternatives, we can decide that the test has good properties in sense of sample size efficiency and power of detection.

4. Some simulation results

In this section we give some simulation results related to the distribution of the plotting statistics and the optimally selected r in various cases. The results of the simulation were based on the S-Plus 4.0. For generation of negative binomial random variates, the S-Plus function rnbinom() was used, and we used the function qnbinom() for the calculation of control lines based on the probability limits.

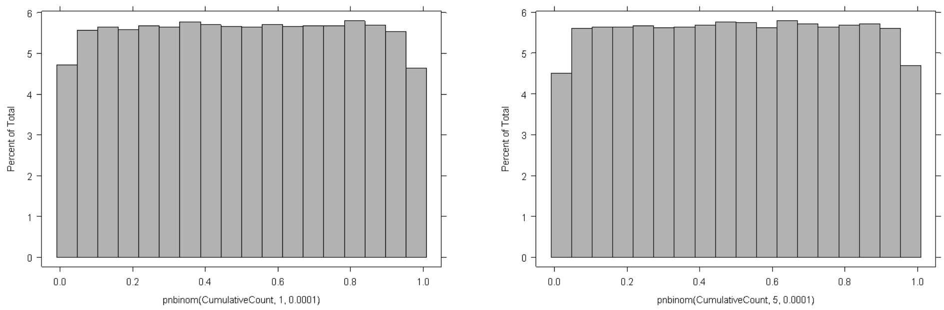
For reference, the control lines obtained by S-Plus in various negative binomial distributions are listed in Table 4.1. We also make various alternatives by delta(δ) in equation of $p_1 = p_0 + \delta p_0$ to investigate the performance of the chart, especially the power of the CPC-r chart. In the equation, if the delta(δ) equals 0 then it means the null situation. And the increasing of delta(δ) means deviation from the null state.

In figure 4.1, the case (a) is shape of negative binomial distribution with r=1 (geometric distribution) and $p = 0.0001$, based on 100,000 random variates. Its expected value is $\frac{1}{p} = 10000$ and it has right skewed distribution. But the case (b) is shape of negative binomial distribution with r=5 and $p = 0.0001$ and the center of distribution moved to $\frac{1}{p} = 10000$. The shape is more similar with the normal distribution than case (a).



(a) r=1 (b) r=5

[Figure 4.1] Shape of negative binomial distribution with $p = 0.0001$, based on 100,000 random variates



(a) r=1 (b) r=5

[Figure 4.2] Shape of distribution of cumulative probability with $p=0.0001$, based on 100,000 random variates

While figure 4.2 shows the distribution of cumulative probability with $p=0.0001$, based on 100,000 random variates. As mentioned previously, it is distributed as uniformly between 0 and 1. In each case (a) r=1 and (b) r=5, the empirical shape of distribution is not so different.

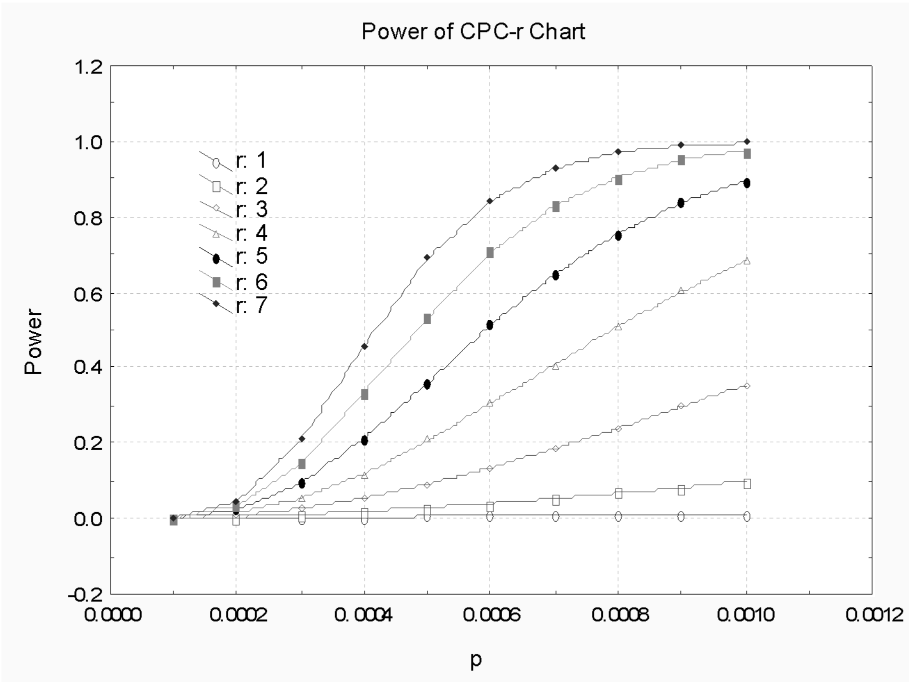
<Table 4.1> Control lines and empirical powers in CPC-r chart

p \ r	$p_0=0.0001(100ppm)$; $p_1=p_0+\delta\cdot p_0$										Quantiles (control lines)
	δ										
	0(α)	1	2	3	4	5	6	7	8	9	0.99865 0.50000 0.00135
1	0.00238 9935 0.0168	0.00218 4992 0.0306	0.00349 3338 0.0732	0.00431 2498 0.1208	0.00606 1999 0.2123	0.00666 1666 0.2798	0.00812 1428 0.3980	0.01009 1251 0.5647	0.01126 1117 0.7059	0.01175 1002 0.8207	66074 6932 14
2	0.00285 20045 0.0100	0.00483 9979 0.0339	0.01124 6659 0.1182	0.01975 5008 0.2761	0.02882 4012 0.5028	0.04011 3335 0.8418	0.05339 2854 1.3096	0.06767 2500 1.8944	0.08208 2223 2.5850	0.09999 1999 3.5017	88999 16784 530
3	0.00277 29977 0.0065	0.00917 15026 0.0427	0.02623 10008 0.1835	0.05479 7508 0.5108	0.09161 5997 1.0693	0.13566 5012 1.8947	0.18542 4282 3.0308	0.23987 3742 4.4876	0.29714 3331 6.2445	0.35493 2995 8.2946	108691 26741 2118
4	0.00265 40011 0.0046	0.01525 19978 0.0534	0.05328 13346 0.2795	0.11794 9996 0.8259	0.20717 8003 1.8122	0.30708 6654 3.2305	0.40843 5709 5.0083	0.51085 4995 7.1587	0.60171 4445 9.4751	0.68374 3997 11.9733	126800 36721 4655
5	0.00253 50048 0.0035	0.02371 24954 0.0665	0.09346 16674 0.3924	0.21411 12473 1.2016	0.36047 10032 2.5152	0.51423 8345 4.3133	0.64840 7134 6.3621	0.75652 6258 8.4627	0.83808 5554 10.5631	0.89613 4991 12.5688	143919 46709 7921
6	0.00248 59876 0.0029	0.03337 29940 0.0780	0.14606 20012 0.5109	0.33426 14967 1.5633	0.53463 12015 3.1148	0.70641 9989 4.9502	0.82960 8560 6.7843	0.90502 7508 8.4383	0.95156 6667 9.9916	0.97634 6004 11.3825	160343 56702 11752
7	0.00276 69997 0.0028	0.04536 35028 0.0906	0.20743 23363 0.6215	0.45913 17492 1.8373	0.68991 13978 3.4550	0.84362 11670 5.0601	0.92891 10018 6.4906	0.97120 8756 7.7646	0.98938 7777 8.9051	0.99565 6997 9.9606	176243 66697 16033

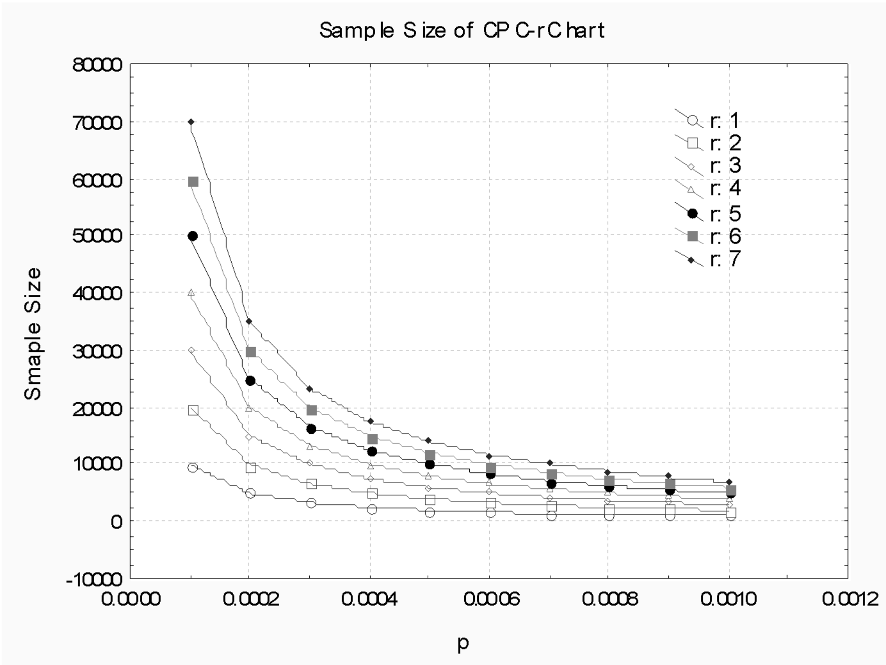
- The first lines mean the power, second lines mean the average sample size, and third lines are (power/size)6000
- In control lines, the first, second and third lines are UCL, CL, LCL, respectively.

In figure 4.3, the performances of CPC-r charts are very low in cases of CPC-1, CPC-2, CPC-3 charts. While in cases of r=4, 5, 6, the performances of CPC-r charts are similar and high. These are distinguished with cases r=1, 2, 3.

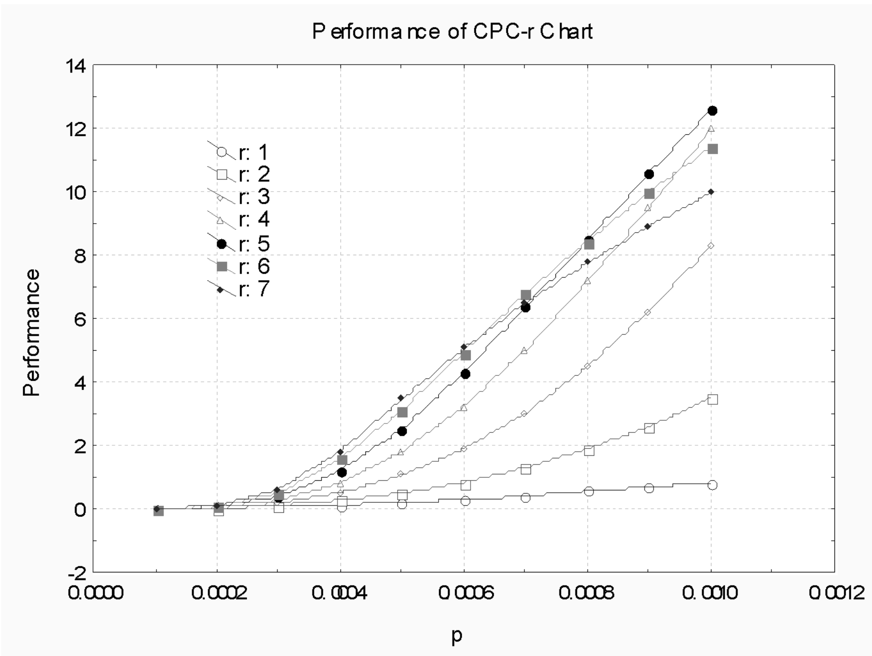
In conclusion, we have a reasonable choice in CPC-r chart as r=4 or 5. In some range of p, the performance of CPC-7 chart is more high than those of others. But the difference is not high and we also know that increase of r causes much more sample size and cost.



(a) Power of CPC-r Chart($p_0 = 100PPM$)



(b) Sample Size of CPC-r Chart($p_0 = 100PPM$)



(c) Performance of CPC-r Chart($p_0 = 100PPM$)

<Figure 4.3> Power, sample size and performance of CPC-r chart

Through this simulation study, we examined other three cases of p_0 $p_0 = 0.001, 0.005, 0.0005$. The results of other 3 cases are very similar with the results of the $p_0 = 0.0001$ explained above. That is, in cases of high yield process with low reject rate the CPC-r chart is very useful and as the optimal value of r 4 or 5 is recommendable.

5. Concluding Remarks

In this paper we discuss some useful methodologies for monitoring of process reject rate p in high yield process. The classical p-chart is not efficient for monitoring of fraction defective, especially in high yield process.

As alternative methods CCC-r or CPC-r chart are considered. Our special concerns are the usefulness of the CPC-r chart, and in this paper we give some simulation results related to the distribution of the plotting statistics in the CPC-r chart and provide a guide line for the selection of r in the CPC-r chart.

References

1. Calvin, T.W.(1983), Quality Control Techniques for Zero Defects, *IEEE Transactions on Component, Hybrids and Manufacturing Technology*, 6, 323-328
2. Chan et al.(2002), Cumulative Probability Control Charts for Geometric and Exponential Process Characteristics, *International Journal of Production Research*, Vol. 40, No. 1, 133-150
3. Statistical Sciences(1994), *S_PLUS for Windows User's Manual*, Siattle : Statistical Sciences.
4. Xie, M and Goh, T.N.(1992), Some Procedures for Decision Making in Controlling High Yield Processes. *Quality and Reliability Engineering International*, 8, 355-360
5. Xie, M and Goh, T.N.(1997), The use of Probability Limits for Process Control based on Geometric Distribution, *International Journal of Quality &Reliability Management,* 14, 64-73
6. Xie, et al.(1998), A Comparative Study of CCC and Cusum Charts, *Quality and Reliability Engineering International*, 14, 339-345

[received date : June 2007, accepted date : July 2007]