$$f_{p}(p) = \frac{(N+1)!}{m!(N-m)!} p^{m} (1-p)^{N-m}$$
(10)

Since binomial distribution assumes more precision than actually exists, it makes control limits more precise than it really is. This problem can be solved using hyperbinomial distribution which considers the distribution of p. If N items are inspected and m of them are found to be good, then the probability of obtaining x good items out of n items is given by hyperbinomial distribution. The corresponding probability mass function (PMF) is as follows:

$$P_{X}(x \text{ of } n \mid m \text{ of } N) = \frac{\binom{m+x}{m}\binom{N+n-m-x}{n-x}}{\binom{N+n+1}{n}}$$
(11)

for x = 0, 1, 2, ..., n, and  $m \le N$ 

The mean and variance of the hyperbinomial distribution is not found in any published literature. However, the mean and variance can be calculated mathematically to be  $\frac{n(m+1)}{(N+2)}$  and

$$\frac{n(n-1)(m+1)(m+2)}{(N+2)(N+3)} + \frac{n(m+1)}{(N+2)} - \left(\frac{n(m+1)}{N+2}\right)^2$$
, respectively (see Appendix). As discussed before, the mean and

variance of the variable sample mean  $\left(\frac{x}{n}\right)$ , is calculated to be  $\frac{(m+1)}{(N+2)}$  and  $\binom{(n-1)(m+1)(m+2)}{(m+1)} \binom{(m+1)^2}{(m+1)^2}$ 

$$\frac{(n-1)(m+1)(m+2)}{n(N+2)(N+3)} + \frac{(m+1)}{n(N+2)} - \left(\frac{m+1}{N+2}\right)$$
, respectively.  
Thus,

$$\mathbf{E}\left(\frac{x}{n}\right) = \frac{(m+1)}{(N+2)} \tag{12}$$

and

$$\operatorname{Var}\left(\frac{x}{n}\right) = \frac{(n-1)(m+1)(m+2)}{n(N+2)(N+3)} + \frac{(m+1)}{n(N+2)} - \left(\frac{m+1}{N+2}\right)^2$$
(13)

Assuming normal approximation, these values of mean and variance are used to construct the control limits of  $3-\sigma$  hyperbinomial control chart as mentioned in (2), (3), and (4).

$$UCL_{HB} = E\left(\frac{x}{n}\right) + k\sqrt{Var\left(\frac{x}{n}\right)} = \frac{(m+1)}{(N+2)} + 3\sqrt{\frac{(m-1)(m+1)(m+2)}{n(N+2)(N+3)}} + \frac{(m+1)}{n(N+2)} - \left(\frac{m+1}{N+2}\right)^2}$$
(14)

$$CL_{HB} = E\left(\frac{x}{n}\right) = \frac{(m+1)}{(N+2)}$$
(15)

$$LCL_{HB} = E\left(\frac{x}{n}\right) - k\sqrt{Var\left(\frac{x}{n}\right)} = \frac{(m+1)}{(N+2)} - 3 \sqrt{\frac{(n-1)(m+1)(m+2)}{n(N+2)(N+3)}} + \frac{(m+1)}{n(N+2)} - \left(\frac{m+1}{N+2}\right)^2$$
(16)

Control charts measure if a process is in statistical control, (i.e., follows normal distribution). Since  $3-\sigma$  encapsulates 99.7% of the data in a normal distribution, if the process falls within that limit, the process is considered to be in statistical control. However, this encapsulation of data can also be done using CDF of the underlying distribution. While using CDF we are interested in fraction conforming instead of fraction non-conforming. The CDF of hyperbinomial distribution can be obtained from its PMF. Suppose, *x* good items are desired with a 99.7% confidence level; this can be obtained using,

$$P(X \ge x) = 1 - P(X < x) = 0.997 \tag{17}$$

The number of x good items that correspond to 99.7% confidence level can then be used to determine whether the process is in statistical control.

#### **3.** Numerical Example

The following example has been taken from Hasin (2007).

San Marino Tube Lights Limited is a famous tube light manufacturing company in North Carolina, producing around 5000 pieces of lights per day. The quality control expert planned to take samples of size 50 units each at every working day. The company worked 22 days in the month under consideration. To test for quality, experts planned to use p-chart. A quality inspector randomly collected and tested 50 tube lights from production line. If they lighted on, they are passed as conforming units, and if they fail, they are rejected as defective units. Data for the 22 working days are shown in Table 1.

Sample No.	No. of	No. of	Fraction	Sample	No. of	No. of	Fraction
<i>(i)</i>	failures	good bulbs	non-	No.	failures	good	non-
	$(x_i)$	$(50-x_i)$	conforming	<i>(i)</i>	$(x_i)$	bulbs	conforming
			$(p_i)$			$(50-x_i)$	$(p_i)$
1	3	47	0.06	12	1	49	0.02
2	2	48	0.04	13	3	47	0.06
3	3	47	0.06	14	2	48	0.04
4	2	48	0.04	15	4	46	0.08
5	3	47	0.06	16	3	47	0.06
6	2	48	0.04	17	3	47	0.06
7	5	45	0.10	18	8	42	0.16
8	3	47	0.06	19	4	46	0.08
9	7	43	0.14	20	2	48	0.04
10	2	48	0.04	21	1	49	0.02
11	1	49	0.02	22	0	50	0.00
Total no. of bulbs $= 1100$			Total failures = 64		Total no. of good bulbs $= 1036$		

Τ	a	bl	le	1	

At first let us construct a control chart based on binomial mean and variance.

Here, Sample size, n = 50. Mean fraction non-conforming,  $E\left(\frac{x}{n}\right) = p = \frac{64}{22 \times 50} = 0.0582$  and variance,

$$\operatorname{Var}\left(\frac{x}{n}\right) = \frac{p(1-p)}{n} = 0.001096.$$

Thus the 3- $\sigma$  control limits of binomial distribution are as follows:

$$UCL_{Binomial} = p + 3\sqrt{\frac{p \times (1-p)}{n}} = 0.0582 + 3\sqrt{\frac{0.0582 \times 0.9418}{50}} = 0.1575$$
$$CL_{Binomial} = p = 0.0582$$
$$LCL_{Binomial} = p - 3\sqrt{\frac{p \times (1-p)}{n}} = 0.0582 - 3\sqrt{\frac{0.0582 \times 0.9418}{50}} \cong 0.00$$

Data is plotted in the control limits calculated above and we obtain the control chart using Minitab shown in Figure 1.

From the control chart, we observe that the proportion of defective items may not be stable. One subgroup (4.5% of the total subgroups) is out of control. So, the process is estimated to be out of control if there is no assignable cause associated with that particular sample (Sample No. 18).

Now we construct a control chart considering that the data follows hyperbinomial distribution. The parameters of hyperbinomial distribution for this numerical problem are as follows.

Total number of inspected items,  $N = 22 \times 50 = 1100$ Total number of defective bulbs from the inspected bulbs, m = 64, lot size, n = 50



Figure 1. p-chart using binomial distribution

Therefore, Mean, 
$$E\left(\frac{x}{n}\right) = \frac{(m+1)}{(N+2)} = 0.0590$$
  
and variance,  $Var\left(\frac{x}{n}\right) = \frac{(n-1)(m+1)(m+2)}{n(N+2)(N+3)} + \frac{(m+1)}{n(N+2)} - \left(\frac{m+1}{N+2}\right)^2$   
= 0.0011594.

Thus, the  $3-\sigma$  control limits for hyperbinomial distribution are calculated as:

$$UCL_{HB} = E\left(\frac{x}{n}\right) + k\sqrt{Var(\frac{x}{n})} = 0.1611$$
$$CL_{HB} = E\left(\frac{x}{n}\right) = 0.0590$$
$$UCL_{HB} = E\left(\frac{x}{n}\right) - k\sqrt{Var(\frac{x}{n})} = -0.0432 \cong 0.$$

Data is plotted in the control limits calculated and we obtain the control chart shown in Fig. 2 using Minitab.



Figure 2. p-chart using hyperbinomial distribution

By observing the control chart we notice that the proportion of defective items is stable. No subgroups are out of control. So we can conclude that the process is in statistical control.

CDF of hyperbinomial distribution can also be used to validate the results. We are interested to determine the least amount of good items (x) that corresponds to 99.7% confidence. As we know, as the confidence level increases, the least amount of good items (x) in the sample decreases. The results are tabulated in Table 2: For the problem under consideration,

Sample size of each inspection, n = 50Total number of inspected bulbs,  $N = 22 \times 50 = 1100$ Total number of non-defective bulbs from the inspected bulbs, M = (1100-64) = 1036

Number of	P(X < x; n = 50, N = 1100,	Number of	P(X < x; n = 50, N = 1100,
good bulbs, x	M = 1036 )	good bulbs, x	M = 1036)
0	5.21E-56	26	1.78E-15
1	2.38E-53	27	1.92E-14
2	5.36E-51	28	1.94E-13
3	7.98E-49	29	1.84E-12
4	8.80E-47	30	1.63E-11
5	7.68E-45	31	1.35E-10
6	5.52E-43	32	1.04E-09
7	3.36E-41	33	7.46E-09
8	1.77E-39	34	4.99E-08
9	8.15E-38	35	3.10E-07
10	3.34E-36	36	1.78E-06
11	1.23E-34	37	9.42E-06
12	4.07E-33	38	4.59E-05
13	1.23E-31	39	2.05E-04
14	3.38E-30	40	8.31E-04
15	8.55E-29	41	3.05E-03
16	2.00E-27	42	1.01E-02
17	4.30E-26	43	2.97E-02
18	8.61E-25	44	7.72E-02
19	1.60E-23	45	1.75E-01
20	2.77E-22	46	3.41E-01
21	4.46E-21	47	5.67E-01
22	6.73E-20	48	7.96E-01
23	9.47E-19	49	9.49E-01
24	1.25E-17	50	1.00E+00
25	1.54E-16		

Table 2

From Table 2, it is observed that at least 42 items are expected to be good with a confidence level of 99.7% considering hyperbinomial distribution. Since all of the samples had at least 42 conforming bulbs, the process is in statistical control. Thus, the result obtained using CDF of hyperbinomial distribution validates the result obtained using hyperbinomial  $3-\sigma$  control chart.

## 4. Conclusion

Binomial distribution is very commonly considered as the underlying distribution of attribute type quality parameters in various product industries. However, when prior estimate of proportion non-conforming is not available, it has to be estimated from limited sample information. Such estimation causes binomial distribution to assume more precision than there actually exists. As a result, it causes the control limits to shrink, which may result in false positive or detection of non-conformity when the process is actually in control, as evident from our example problem. Binomial control charts also fail to utilize additional sample information. Thus, this study proposes the use of hyperbinomial distribution for the construction of  $3-\sigma$  control chart when limited sample information is available. Hyperbinomial distribution considers the variability in proportion non-conforming. This reduces the rate of false detection significantly as shown in the numerical problem. The study also validates the use of hyperbinomial  $3-\sigma$  control chart by comparing with the result obtained using CDF of hyperbinomial distribution. This efficient approach can be used for quality inspection in various manufacturing industries. In future, normal approximation to hyperbinomial distribution will be evaluated as a part of this study since it serves as the basis of construction of Shewhart control charts and will be applied to other numerical problems.

## Acknowledgement

The author of this paper is grateful to the Department of Industrial and Production Engineering (IPE), BUET for providing sufficient computational facility to carry out the research.

## References

- Hasin, A. A. (2007). Quality control and management. Bangladesh Business Solutions, Dhaka.
- Montgomery, Douglas C. Introduction to statistical quality control. John Wiley & Sons, 2007.
- Das, N. (2008). Non-parametric control chart for controlling variability based on rank test. *Economic Quality Control*, 23(2), 227-242.
- Sellers, K. F. (2012). A generalized statistical control chart for over-or under-dispersed data. *Quality and Reliability Engineering International*, 28(1), 59-65.
- Feller, W. (2008). An introduction to probability theory and its applications (Vol. 2). John Wiley & Sons.
- Cheng, S. W. and Xie, H. (2000). Control Charts for Lognormal data, *Tamkang Journal of Science and Engineering*, Vol. 3, No. 3, 131-137
- Chakraborti, S., Van der Laan, P., & Bakir, S. T. (2001). Nonparametric control charts: an overview and some results. *Journal of Quality Technology*, 33(3), 304.
- Bakir, S. T. (2001, August). Classification of distribution-free quality control charts. In Proceedings of the Annual Meeting of the American Statistical Association (pp. 5-9).
- Park, C., Park, C., Reynolds Jr, M. R., & Reynolds Jr, M. R. (1987). Nonparametric procedures for monitoring a location parameter based on linear placement statistics. *Sequential Analysis*, 6(4), 303-323.
- Orban, J., & Wolfe, D. A. (1982). A class of distribution-free two-sample tests based on placements. *Journal of the American Statistical Association*, 77(379), 666-672.
- Shmueli, G., Minka, T. P., Kadane, J. B., Borle, S., & Boatwright, P. (2005). A useful distribution for fitting discrete data: revival of the Conway–Maxwell–Poisson distribution. *Journal of the Royal Statistical Society: Series C* (Applied Statistics), 54(1), 127-142.
- Jackson, J. E. (1972). All count distributions are not alike. *Journal of Quality Technology*, Vol 4, No 2, P 86-92, April 1972. 4 Tab, 15 Ref.
- Kaminsky, F. C., Benneyan, J. C., Davis, R. D., & Burke, R. J. (1992). Statistical control charts based on a geometric distribution. *Journal of Quality Technology*, 24(2).
- Joekes, S., & Barbosa, E. P. (2013). An improved attribute control chart for monitoring non-conforming proportion in high quality processes. *Control Engineering Practice*, 21(4), 407-412.
- Fisher, S. R. A., & Cornish, E. A. (1960). The percentile points of distributions having known cumulants. *Technometrics*, 2(2), 209-225.
- Haldar, A., & Mahadevan, S. (2000). Probability, reliability, and statistical methods in engineering design. John Wiley.
- Haldar, A. (1982). Statistical and Probabilistic Methods in Geomechanics. In Numerical Methods in *Geomechanics* (pp. 473-504). Springer, Dordrecht

# Appendix

Derivation of Mean and Variance for Hyperbinomial Distribution: We know, the PMF of hyperbinomial distribution is

$$P_X\left(x \text{ of } n \mid m \text{ of } N\right) = \frac{\binom{m+x}{m}\binom{N+n-m-x}{n-x}}{\binom{N+n+1}{n}}, \text{ for } x = 0, 1, 2, \dots, n, \text{ and } m \le N$$
  
Since,  $\binom{p}{r} = \binom{p}{p-r}, \text{ so } \binom{m+x}{m} = \binom{m+x}{x}, \text{ we can write,}$ 
$$P_X\left(x \text{ of } n \mid m \text{ of } N\right) = \frac{\binom{m+x}{x}\binom{N+n-m-x}{n-x}}{\binom{N+n+1}{n}}, \text{ for } x = 0, 1, 2, \dots, n, \text{ and } m \le N$$

Now,

$$x\binom{m+x}{x} = \frac{x(m+x)!}{x!m!}$$
$$= \frac{x(m+x)!}{x(x-1)!m!} \times \frac{(m+1)}{(m+1)}$$

$$= \frac{(m+1)[(m+1)+(x-1)]!}{(x-1)!(m+1)!}$$
$$= (m+1) \frac{(m+1)+(x-1)}{(m+1)}$$

Again,

$$\binom{N+n+1}{n} = \frac{\binom{N+n+1}{n!}}{\binom{N+1}{n!}}$$
$$= \frac{\left[\binom{N+1}{n!} + \binom{n-1}{n!} + 1\right]!}{\binom{N+1}{n!}} \times \frac{\binom{N+2}{N+2}}{\binom{N+2}{n!}}$$
$$= \frac{\binom{N+2}{n}}{\binom{N+1}{n!}} \binom{\binom{N+1}{n} + \binom{n-1}{n!}}{\binom{N-1}{n!}}$$

Mean of hyperbinomial distribution can be found as follows:

Since the expression inside the summation symbol is analogous to the PMF of hyperbinomial distribution, the summation over the full range of x must be equal to 1. So we can write:

$$\mu_{\rm X} = \frac{n(m+1)}{(N+2)}$$

Variance of hyperbinomial distribution can be found as follows:

$$Var(X) = E(X^{2}) - [E(X)]^{2}$$
  
=  $E[X(X-1)+X] - [E(X)]^{2}$   
=  $E[X(X-1)] + E[X] - [E(X)]^{2}$   
=  $E[X(X-1)] + E[X](1 - [E(X)])$   
=  $\sum_{x=0}^{h} x(x-1)P_{X}(x \text{ of } n \mid m \text{ of } N) + \mu_{X}(1-\mu_{X})$   
=  $\sum_{x=0}^{h} \frac{x(x-1)\binom{m+x}{x}\binom{N+n-m-x}{n-x}}{\binom{N+n+1}{n}}$   
+  $\mu_{X}(1-\mu_{X})$ 

Now,

$$\begin{aligned} x(x-1)\binom{m+x}{x} &= \frac{x(x-1)(m+x)!}{m!x!} \\ &= \frac{x(x-1)(m+x)!}{x(x-1)(x-2)!m!} \times \frac{(m+1)(m+2)}{(m+1)(m+2)} \\ &= \frac{\left[(m+2)+(x-2)\right]!}{(x-2)!(m+2)!} \times (m+1)(m+2) \\ &= (m+1)(m+2)\binom{(m+2)+(x-2)}{(m+2)} \end{aligned}$$

Again,

$$\binom{N+n+1}{n} = \frac{(N+n+1)!}{n!(N+1)!}$$
  
=  $\frac{\left[(N+2)+(n-2)+1\right]!}{n(n-1)(n-2)!(N+1)!} \times \frac{(N+2)(N+3)}{(N+2)(N+3)}$   
=  $\frac{(N+2)(N+3)}{n(n-1)} \times \frac{\left[(N+2)+(n-2)+1\right]!}{(n-2)!(N+3)!}$   
=  $\frac{(N+2)(N+3)}{n(n-1)} \binom{(N+2)+(n-2)+1}{n-2}$   
 $\therefore \operatorname{Var}(X) = \sum_{x=0}^{h} \frac{x(x-1)\binom{m+x}{x}\binom{N+n-m-x}{n-x}}{\binom{N+n+1}{n}}$   
 $+\mu_{X}(1-\mu_{X})$ 

Since the term

$$\frac{\left[\binom{(m+2)+(x-2)}{(m+2)}\right]\binom{(N+2)+(n-2)-(m+2)-(x-2)}{(n-2)-(x-2)}}{\left[\binom{(N+2)+(n-2)+1}{n-2}\right]}$$

is analogous to the PMF of hyperbinomial distribution, summation of this term over full range of x must be equal to 1. Hence, ( $x = x^2 + x^2 +$ 

$$\operatorname{Var}(X) = \frac{n(n-1)(m+1)(m+2)}{(N+2)(N+3)} \times 1 + \mu_{X}(1-\mu_{X})$$
$$= \frac{n(n-1)(m+1)(m+2)}{(N+2)(N+3)} + \frac{n(m+1)}{(N+2)} \left[ 1 - \frac{n(m+1)}{(N+2)} \right]$$
$$= \frac{n(m+1)[N^{2} + Nm - nm + nN - 2n + 3N - 2m + 2]}{(N+2)^{2}(N+3)}$$

$$= (m+1)(m+2) \binom{(m+2)+(x-2)}{(m+2)} \times \sum_{x=2}^{h} \frac{\binom{(N+2)+(n-2)-(m+2)-(x-2)}{(n-2)-(x-2)}}{n(n-1)} + \mu_X (1-\mu_X)$$
$$= \frac{n(n-1)(m+1)(m+2)}{(N+2)(N+3)} \times \sum_{x=2}^{h} \frac{\binom{(m+2)+(x-2)}{(m+2)} \binom{(N+2)+(n-2)+(n-2)-(x-2)}{(n-2)-(x-2)}}{\binom{(N+2)+(n-2)+1}{n-2}} + \mu_X (1-\mu_X)$$

Now,

$$x(x-1)\binom{m+x}{x} = \frac{x(x-1)(m+x)!}{m!x!}$$
$$= \frac{x(x-1)(m+x)!}{x(x-1)(x-2)!m!} \times \frac{(m+1)(m+2)}{(m+1)(m+2)}$$
$$= \frac{[(m+2)+(x-2)]!}{(x-2)!(m+2)!} \times (m+1)(m+2)$$
$$= (m+1)(m+2)\binom{(m+2)+(x-2)}{(m+2)}$$

$$\binom{N+n+1}{n} = \frac{(N+n+1)!}{n!(N+1)!}$$
$$= \frac{[(N+2)+(n-2)+1]!}{n(n-1)(n-2)!(N+1)!} \times \frac{(N+2)(N+3)}{(N+2)(N+3)}$$
$$= \frac{(N+2)(N+3)}{n(n-1)} \times \frac{[(N+2)+(n-2)+1]!}{(n-2)!(N+3)!}$$
$$= \frac{(N+2)(N+3)}{n(n-1)} \binom{(N+2)+(n-2)+1}{n-2}$$

$$\therefore \operatorname{Var}(X) = \sum_{x=0}^{h} \frac{x(x-1)\binom{m+x}{x}\binom{N+n-m-x}{n-x}}{\binom{N+n+1}{n}} + \mu_{x}(1-\mu_{x})$$

$$= \sum_{x=2}^{h} \frac{(m+1)(m+2)\binom{(m+2)+(x-2)}{(m+2)}\binom{(N+2)+(n-2)-(m+2)-(x-2)}{(n-2)-(x-2)}}{\frac{(N+2)(N+3)}{n(n-1)}\binom{(N+2)+(n-2)+1}{n-2}} + \mu_{x}(1-\mu_{x})$$

$$= \frac{n(n-1)(m+1)(m+2)}{(N+2)(N+3)} \sum_{x=2}^{h} \frac{\binom{(m+2)+(x-2)}{(m+2)}\binom{(N+2)+(n-2)-(m+2)-(x-2)}{(n-2)-(x-2)}}{\binom{(N+2)+(n-2)+1}{n-2}} + \mu_{x}(1-\mu_{x})$$

Since the term

$$\frac{\left[\binom{(m+2)+(x-2)}{(m+2)}\right]^{(N+2)+(n-2)-(m+2)-(x-2)}}{\binom{(n+2)+(n-2)+1}{n-2}}$$

is analogous to the PMF of hyperbinomial distribution, summation of this term over full range of x must be equal to 1. Hence,

$$\operatorname{Var}(X) = \frac{n(n-1)(m+1)(m+2)}{(N+2)(N+3)} \times 1 + \mu_{X}(1-\mu_{X})$$
$$= \frac{n(n-1)(m+1)(m+2)}{(N+2)(N+3)} + \frac{n(m+1)}{(N+2)} \left[ 1 - \frac{n(m+1)}{(N+2)} \right]$$
$$= \frac{n(m+1)[N^{2} + Nm - nm + nN - 2n + 3N - 2m + 2]}{(N+2)^{2}(N+3)}$$

#### **Biography / Biographies**

**Shourav Ahmed** is a Lecturer in the Department of Industrial and Production Engineering (IPE) at Bangladesh University of Engineering and Technology (BUET), Dhaka, Bangladesh. He earned his B.Sc. and M.Sc. degrees in Industrial and Production Engineering from BUET, Dhaka, Bangladesh in 2015 and 2018, respectively. He has completed research projects with Nestle, National Polymer and Walton. His research interests include quality management, supply chain management, micro/nano fabrication, multidisciplinary design optimization, uncertainty quantification, etc. He is an esteemed member of IEB, ORCA, and AIPE.

**Gulam Kibria** is a Lecturer in the Department of IPE at BUET, Dhaka, Bangladesh. He earned his B.Sc. and M.Sc. degrees in Industrial and Production Engineering from BUET, Dhaka, Bangladesh in 2015 and 2018, respectively. His research interests include operations management, operations research and decision analysis, multidisciplinary system analysis, uncertainty and risk management, etc.

**Kais Zaman** is a professor in the Department of Industrial and Production Engineering (IPE) at Bangladesh University of Engineering and Technology (BUET), Dhaka, Bangladesh. Currently, he is also the head of the department. He earned his B.Sc. and M.Sc. degrees in Industrial and Production Engineering from BUET, Dhaka, Bangladesh in 2003 and 2005, respectively. He also earned his M.Sc. and Ph.D. degrees in Reliability and Risk Engineering and Management from Vanderbilt University in 2010. His research interests include artificial intelligence and machine learning, operations research and decision analysis, uncertainty and risk management, etc. He has publications in numerous prestigious journals and conferences.