Proceedings of the World Congress on Engineering and Computer Science 2011 Vol II
WCECS 2011, October 19-21, 2011, San Francisco, USA

D Chart: An Efficient Alternative to Monitor Process Dispersion

Saddam Akber Abbasi and Arden Miller *

Abstract—**Control chart is the most important Statistical Process Control tool used to monitor reliability and performance of industrial processes. For monitoring changes in process dispersion, the R and S charts are widely used. These control charts perform better under the ideal assumption of normality but are well known to be very inefficient in presence of outliers or departures from normality. In this study we propose a new control chart for monitoring process dispersion, namely the D chart, and compared its performance with the R and S charts using probability to signal as a performance measure. It has been observed that the newly proposed chart is superior to the R chart and is a close competitor to the S chart under normality of quality characteristic. When the assumption of normality is violated, the D chart is more powerful than both the R and S charts in terms of detecting shifts in process dispersion. This study will help quality practitioners to choose an efficient alternative to the classical R and S charts for monitoring dispersion of industrial processes.**

Keywords:* *Process Dispersion, Variability Charts, Monte Carlo Simulations, Probability to Signal, Non-Normality

1 Introduction

Control chart introduced by Walter A. Shewhart in 1920's, is the most important Statistical Process Control (SPC) tool used to monitor reliability and performance of industrial processes. The basic purpose of implementing control chart procedures is to detect abnormal variations in the process (location & scale) parameters. Although first proposed for manufacturing industry, control charts have recently been applied in a wide variety of disciplines, such as in nuclear engineering ([8]), health care ([18]), education ([17]) , analytical laboratories ([1]) etc.

Monitoring process dispersion is an important component of SPC. Dispersion control charts are a well known tool used for improving process capability and productivity by reducing variability in the process. The R and S charts are the two most widely used control charts for monitoring changes in process dispersion ([11]). The design of these charts is based on estimating the process standard deviation σ using sample range and sample standard deviation respectively. These charts perform better under the ideal assumptions but are well known to be very inefficient when the assumption of normality is violated. In this study we propose a new dispersion control chart, namely the D chart, based on Downton's based estimate of process standard deviation. The design of the D chart is established and is shown to be more efficient as compared to the classical R and S charts, particularly for non-normal processes.

Assume X be a normally distributed quality characteristic with in-control mean μ and standard deviation σ (i.e. $X \sim N(\mu, \sigma^2)$) . Let $X_1, X_2, \cdots, X_n$ represents a random sample of size n and $X_{(1)}, X_{(2)}, \cdots, X_{(n)}$ be the corresponding order statistics. The Downton's estimator is defined as (see [5] and [2]):

$$D = \frac{2\sqrt{\pi}}{n(n-1)} \sum_{i=1}^{n} \left[i - \frac{1}{2}(n+1) \right] X_{(i)} \qquad (1)$$

For normally distributed quality characteristic, D is an unbiased estimator of σ (see [3]) and it has been shown in the past that D is not much affected by non-normality. The purpose of this study is to develop a variability chart based on D that performs better than existing variability charts under the existence and violation of normality assumption. The rest of this study is organized as follows: In the next section the widely used 3-sigma and probability limit structure of the D chart is established following [16] and [7]. The following section compares the performance of the D, R and S charts assuming normality of quality characteristics. The comparison is made using probability to signal as a performance measure. Fourth section presents comparison of these charts when the assumption of normality is violated and quality characteristic is assumed to follow non-normal (heavy tailed symmetric and skewed) distributions following [16] and [13]. Finally conclusions have been made in the last section.

2 Design of D Control Chart

Suppose the relationship between D and σ be defined by a random variable Z as $Z = D/\sigma$ (similar to $W = R/\sigma$ for the R chart – see [11]). For setting up control limits of

*Department of Statistics, The University of Auckland, New Zealand. Corresponding author: Sadddam Akber Abbasi, Email: sabb025@aucklanduni.ac.nz

ISBN: 978-988-19251-7-6
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

the proposed D chart, estimates of σ and σ_D are required. By taking expectations on both sides of Z, we obtain:

$$E(Z) = E(D/\sigma) = E(D)/\sigma \tag{2}$$

$E(D)$ can be replaced with average of sample $D's$ ($\overline{D}$), computed from an appropriate number of random samples obtained from a process during normal operating conditions (similar to $\overline{R}$ and $\overline{S}$ used in the construction of R and S charts). Let $E(Z) = z_2$, as D is an unbiased estimator of σ hence we have $z_2 = 1$ (for every value of n). Thus under normality, an unbiased estimator of σ based on Downton's estimator is given as $\widehat{\sigma} = \overline{D}$. Similarly for an estimate of σ_D we have $\sigma_Z = \sigma_D/\sigma$. Let $\sigma_Z = z_3$, hence we have

$$\sigma_D = z_3\sigma \tag{3}$$

Ref. [3] showed that

$$var(D) = \frac{\sigma^2}{n(n-1)}\left\{n(\frac{1}{3}\pi + 2\sqrt{3} - 4) + (6 - 4\sqrt{3} + \frac{1}{3}\pi)\right\} \tag{4}$$

From Equations (3) and (4) we have

$$z_3 = \frac{1}{\sqrt{n(n-1)}}\sqrt{n(\frac{1}{3}\pi + 2\sqrt{3} - 4) + (6 - 4\sqrt{3} + \frac{1}{3}\pi)} \tag{5}$$

Replacing an estimate of σ (i.e. $\widehat{\sigma} = \overline{D}$) in Equation (3), we obtain $\widehat{\sigma}_D = z_3\overline{D}$
Hence the widely used 3-sigma control limits for the proposed D chart are defined as

$$\begin{aligned} LCL &= max(0, \overline{D} - 3z_3\overline{D}) = Z_3\overline{D} \\ CL &= \overline{D} \\ UCL &= \overline{D} + 3z_3\overline{D} = Z_4\overline{D} \end{aligned} \tag{6}$$

where $Z_3 = max(0, 1-3z_3)$ and $Z_4 = 1+3z_3$. We can observe that the coefficients z_3, Z_3 and Z_4 entirely depends upon the sample size n. After setting up control limits, sample statistic D is plotted against time or sample number. If all the $D's$ lie inside control limits we can say that the process variability is in statistical control otherwise if one or more $D's$ lie outside control limits, the process variability is said to be out-of-control.

The use of 3-sigma limits is based on the symmetric assumption of the plotted statistic, we will see that the distribution of D is not symmetric atleast for small to moderate values of n. Hence there is a need to develop the probability limit structure for the proposed D chart. Probability limits for the D chart can be computed by using quantile points of the distribution of Z. Let α be the specified probability of making Type-I error, denoting α-quantile of the distribution of Z by Z_α, the probability limits based on D are given as:

$$\begin{aligned} LCL &= Z_{(\alpha/2)}\widehat{D} \quad \text{with} \quad \Pr(Z \leq Z_{(\alpha/2)}) = \alpha/2 \\ UCL &= Z_{(1-\alpha/2)}\widehat{D} \quad \text{with} \quad \Pr(Z \geq Z_{(1-\alpha/2)}) = 1 - \alpha/2 \end{aligned} \tag{7}$$

These quantile points have been computed through extensive Monte Carlo simulation routines. The distribution of Z is obtained by generating 10,000 samples of size $n = 2, 3, \cdots, 15, 20, 25, 35, 50, 75$ and 100 from standard normal distribution. For a specified Type-I error probability α, $(\alpha/2)^{th}$ and $((1-\alpha)/2)^{th}$ quantile points have been computed from the distribution of Z for every combination of α and n. The same procedure is repeated 1000 times and the mean values of the quantile points together with their standard errors (in parenthesis) are reported in Table 1 (for parent normal distribution). The 3-sigma and probability limit structure of R and S charts with their respective control chart constants and quantile points can be seen in [15].

3 Comparison of D, R and S Charts for Normal Processes

The probability to signal shifts in the process dispersion is used as a performance measure following [6, 12] and [14]. In our case, the process is said to be out-of-control whenever process standard deviation σ shifts from an in-control value, say σ_0 to another value say σ_1, where σ_1 is defined as $\sigma_1 = \sigma_0 + \delta\sigma_0$. For a fixed false alarm rate, control chart structure which gives highest probability to signal for out-of-control situations will indicate best performance as compared to other charts.

After setting up the probability limits for $\alpha = 0.002$, probability to signal have been computed for both in-control and out-of-control situations for the D, R and S charts using their respective control chart coefficients and quantile points. To save space and to aid in visual clarity, power curves have been constructed instead of presenting results in tabular form. The power curves of the three charts for normally distributed quality characteristics for $n = 5, 10$ and 15 are shown in Figure 1.

From power curves in Figure 1 we can observe that for zero sigma shift in process standard deviation, the probability of signaling is very close to 0.002 for all the charts and for every sample size, representing the case for an in-control process. When the process is out-of-control, the D chart is equally efficient to the S chart for detecting shifts in process variability and have significantly higher probability to signal as compared to the R chart, as the power curves of the D chart coincides with that of the S chart and remains always higher than the power curves of the R chart for every choice of n. Hence we can say that under the ideal assumption of normality, the D chart is more efficient than the R chart and acts as a close competitor to the S chart.

ISBN: 978-988-19251-7-6
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

Table 1: Quantile points of the distribution of Z

n	$z_{0.001}$	$z_{0.01}$	$z_{0.05}$	$z_{0.95}$	$z_{0.99}$	$z_{0.999}$
2	0.00166	0.01565	0.0785	2.45975	3.23694	4.15268
	(0.00056)	(0.00015)	(0.001)	(0.00027)	(0.00111)	(0.00033)
3	0.03551	0.11492	0.25502	1.96152	2.439	3.00633
	(0.00084)	(0.00057)	(0.00086)	(0.00052)	(8e-04)	(0.00086)
4	0.09823	0.21368	0.37075	1.76101	2.12345	2.53082
	(0.00032)	(0.00111)	(0.00031)	(0.00109)	(0.00096)	(0.00069)
5	0.15217	0.28511	0.44493	1.64577	1.9552	2.30548
	(0.00025)	(0.00044)	(0.00027)	(0.00047)	(0.00076)	(0.00089)
6	0.20846	0.34765	0.50023	1.56683	1.8354	2.15353
	(0.00025)	(0.00048)	(0.00032)	(0.00057)	(0.00093)	(0.00036)
7	0.25049	0.39398	0.54303	1.51488	1.75368	2.02625
	(0.00013)	(2e-04)	(0.00061)	(0.00034)	(0.00076)	(0.00035)
8	0.30186	0.43455	0.57671	1.47484	1.68721	1.94266
	(0.00033)	(0.00087)	(0.00023)	(0.00025)	(0.00086)	(0.00028)
9	0.33462	0.46845	0.60199	1.443	1.64455	1.87114
	(0.00016)	(0.00025)	(0.00079)	(0.00011)	(0.00035)	(0.00081)
10	0.37155	0.49531	0.62393	1.41335	1.60123	1.81688
	(0.00028)	(0.00024)	(0.00046)	(0.00078)	(0.00063)	(0.00035)
11	0.38631	0.51896	0.63972	1.39357	1.56754	1.7785
	(9e-04)	(0.00036)	(0.00091)	(0.00024)	(0.00068)	(0.00073)
12	0.41242	0.53575	0.6575	1.37317	1.54059	1.7294
	(0.00109)	(0.00087)	(0.00039)	(0.00056)	(0.00076)	(0.00085)
13	0.43215	0.55201	0.6712	1.35789	1.52125	1.69148
	(0.00067)	(0.00034)	(0.00079)	(0.00021)	(0.00025)	(0.00063)
14	0.45467	0.57071	0.68466	1.34069	1.49412	1.68263
	(0.00031)	(0.00056)	(0.00077)	(0.00041)	(0.00033)	(0.00039)
15	0.46773	0.58356	0.6935	1.32709	1.47725	1.63709
	(0.00066)	(0.0011)	(0.00098)	(0.00022)	(0.00051)	(0.00081)
20	0.54297	0.63953	0.73694	1.27848	1.4035	1.54339
	(0.00042)	(0.00033)	(0.00092)	(0.00059)	(0.00017)	(0.00011)
25	0.57977	0.67448	0.76569	1.24488	1.35738	1.47917
	(0.00074)	(0.00028)	(0.00107)	(0.00053)	(0.00088)	(0.00072)
35	0.64444	0.72547	0.8022	1.20638	1.29519	1.4015
	(0.00022)	(0.00091)	(0.00012)	(0.00031)	(0.00067)	(0.00079)
50	0.70201	0.77009	0.83552	1.17129	1.24661	1.33503
	(7e-04)	(0.00084)	(0.00019)	(5e-04)	(0.00052)	(0.00094)
75	0.75708	0.8122	0.86561	1.13826	1.1971	1.26395
	(0.00031)	(0.00099)	(0.00053)	(0.00036)	(0.00097)	(0.00025)
100	0.78607	0.83591	0.88294	1.12025	1.17153	1.2276
	(0.00094)	(0.00079)	(0.00055)	(0.00105)	(0.00077)	(0.0011)

4 Comparison of D, R and S Charts for Non-Normal Processes

Normal distribution have wide applications in statistics and almost all SPC charts are based on this assumption. But in practice data from many real world processes follow non-normal distributions. To mention a few of such cases: [4] and [9] pointed out that quality characteristics such as capacitance, insulation resistance, surface finish, roundness, mold dimensions follow non-normal distributions. [10] indicates that impurity levels in semiconductor process chemicals follow Gamma distribution. Many other characteristics such as straightness, flatness, cycle time are not distributed normally. Hence there is a need to study the performance of these variability charts for different parent non-normal distributions. To represent the case of non-normal processes, the performance of D, R and S charts is investigated by assuming that the quality characteristic follows heavy tailed symmetric (Student's t) and skewed (Gamma and Weibull) distributions. The density function of these non-normal distributions are given as:

Student's t (t_k):

$$f(x|k) = \frac{\Gamma[(k+1)/2]}{\sqrt{k\pi}\Gamma(k/2)}\left(1+\frac{x^2}{k}\right)^{-(k+1)/2},$$

$$-\infty < x < \infty, k > 0$$

Gamma(α, β):

$$f(x|\alpha,\beta) = \frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x},$$

$$x > 0, \alpha > 0, \beta > 0$$

Weibull(α, β):

$$f(x|\alpha,\beta) = \frac{\alpha}{\beta}x^{\alpha-1}e^{-x^\alpha/\beta},$$

$$x \geq 0, \alpha > 0, \beta > 0$$

Probability to signal of the D, R and S charts have been computed for these non-normal distributions using similar simulation routines as were used earlier for the case of normal distribution. In our simulation study we used Student's t distribution with $k = 5$, Gamma distribution with $\alpha = 2$ and $\beta = 1$, and finally Weibull distribution with $\alpha = 1.5$ and $\beta = 1$. The power curves of the three charts when quality characteristic is assumed to follow Student's t, Gamma and Weibull distributions are presented in Figures 2-4 respectively.

From Figures 2-4 we can clearly see that the power curves of D chart are always higher than the power curves of both the R and S charts for all non-normal cases and for every choice of sample size n. This indicates that the D chart has higher probability to signal shifts in process variability as compared to both R and S charts when the assumption of normality is violated. We can also observe that the difference in the detection ability of these charts increases with an increase in n. Relatively the R chart is extremely affected while the D chart is least affected by non-normality. Hence for non-normal processes, we can easily say that the proposed D chart is always superior than both the classical R and S charts.

5 Conclusions

This study proposes an efficient control chart, namely the D chart, to monitor process dispersion. The performance of the D chart is compared with the classical R and S charts using probability to signal as a performance measure. It has been shown that for normally distributed quality characteristic, the D chart is equally efficient to the S chart in terms of detecting shifts in process variability and has significantly better detection ability as compared to the R chart. For non-normal processes, the D chart clearly showed superiority over both the R and S charts. Quality control practitioners can now easily choose the proposed D chart as a superior alternative to both the classical R and S charts due to its efficient detection ability.

References

[1] S.A. Abbasi, *On the performance of EWMA chart in the presence of two-component measurement error*, Quality Engineering 22 (2010), pp. 199–213.

[2] M.O. Abu-Shawiesh and M.B. Abdullah, *Estimating the process standard deviation based on downton's*

estimator, Quality Engineering 12 (2000), pp. 357–363.

[3] F.C. Barnett, K. Mullen, and J.G. Saw, *Linear estimates of a population scale parameter*, Biometrika 54 (1967), pp. 551–554.

[4] D. Bissell, 1st ed., Chapman & Hall, New York 1994.

[5] F. Downton, *Linear estimates with polynomial coefficients*, Biometrika 53 (1966), pp. 129–141.

[6] A.J. Duncan, *Operating characteristics of R charts*, Industrial Quality Control 7 (1951), pp. 40–41.

[7] I.M. Gonzalez and E. Viles, *Design of R control chart assuming a gamma distribution*, Economic Quality Control 16 (2001), pp. 199–204.

[8] S.L. Hwang, J.T. Lin, G.F. Liang, Y.J. Yau, T.C. Yenn, and C.C. Hsu, *Application control chart concepts of designing a pre-alarm system in the nuclear power plant control room*, Nuclear Engineering and Design 238 (2008), pp. 3522–3527.

[9] P.C. James, *C_pk equivalencies*, Quality 28 (1989), p. 75.

[10] W.A. Levinson and A. Polny, *Spc for tool particle counts*, Semiconductor International 22 (June, 1999), pp. 117–121.

[11] D.C. Montgomery, Introduction to statistical quality control, 4rth ed., Wiley, New York, 2001.

[12] P.R. Nelson, *Power curves for the analysis of means*, Technometrics 27 (1985), p. 6573.

[13] M.F. Ramalhoto and M. Morais, *Shewhart control charts for the scale parameter of a weibull control variable with fixed and variable sampling intervals*, Journal of Applied Statistics 26 (1999), pp. 1129–160.

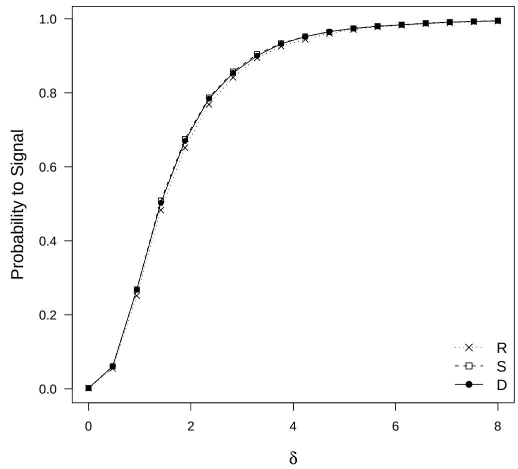
[14] M. Riaz, *A dispersion control chart*, Communications in Statistics: Simulation and Computation 37 (2008), pp. 1239–1261.

[15] P.R. Ryan, Statistical methods for quality improvement, 2nd ed., Wiley, New York, 2000.

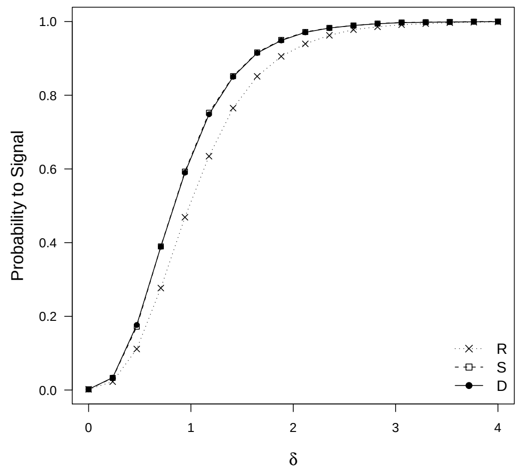
[16] W.A. Shewhart, D. Van Nostrand; reprinted by the American Society for Quality Control in 1980, Milwauker,WI., New York 1931.

[17] Z. Wang and R. Liang, *Discuss on applying SPC to quality management in university education*, in *Proceedings of the 9th International Conference for Young Computer Scientists, ICYCS 2008*, 2008, pp. 2372–2375.

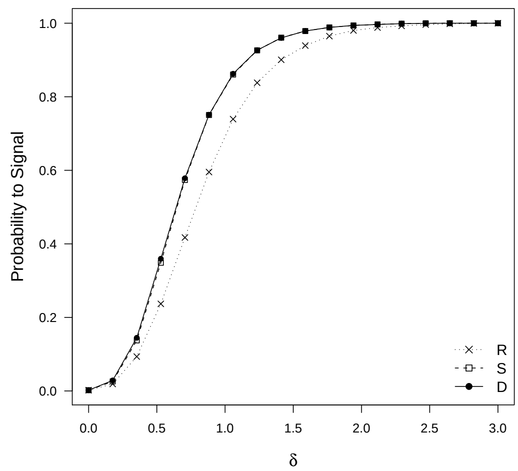
[18] W.H. Woodall, *The use of control charts in healthcare and public-health surveillance*, Journal of Quality Technology 38 (2006), pp. 89–104.



(a) $n = 5$



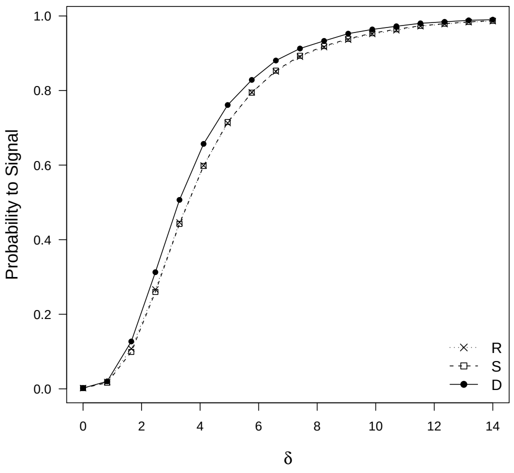
(b) $n = 10$



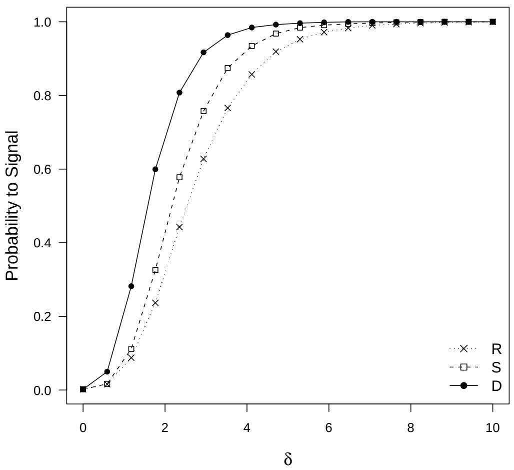
(c) $n = 15$

Figure 1: Power curves of D, R and S charts for $n = 5, 10$ and 15 under Normal distribution when $\alpha = 0.002$

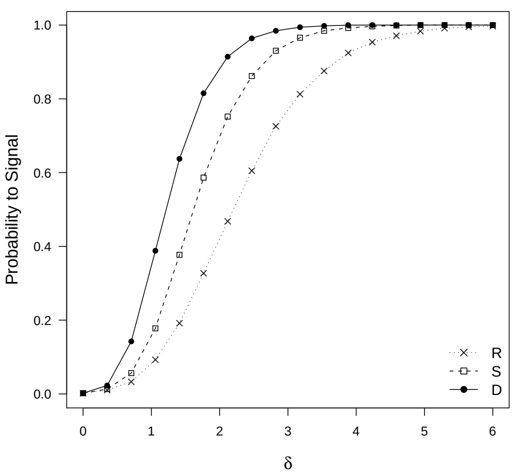
ISBN: 978-988-19251-7-6
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)



(a) $n = 5$

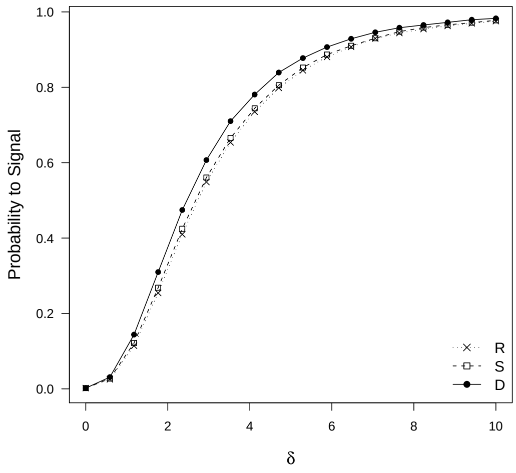


(b) $n = 10$

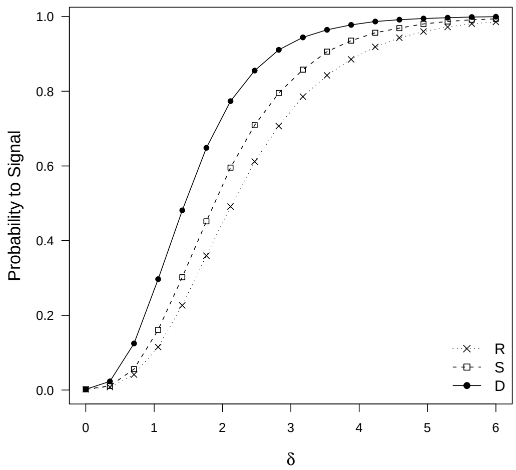


(c) $n = 15$

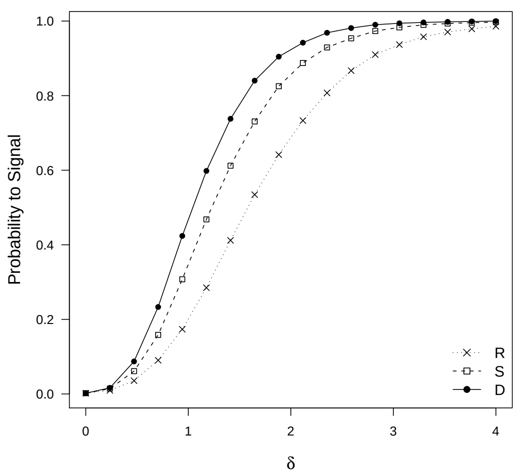
Figure 2: Power curves of D, R and S charts for $n = 5, 10$ and 15 under Student's t distribution when $\alpha = 0.002$



(a) $n = 5$



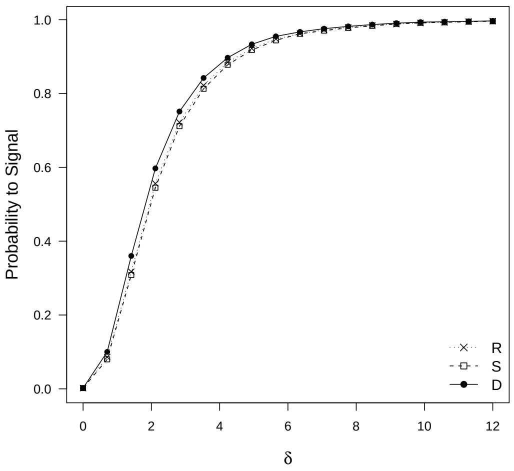
(b) $n = 10$



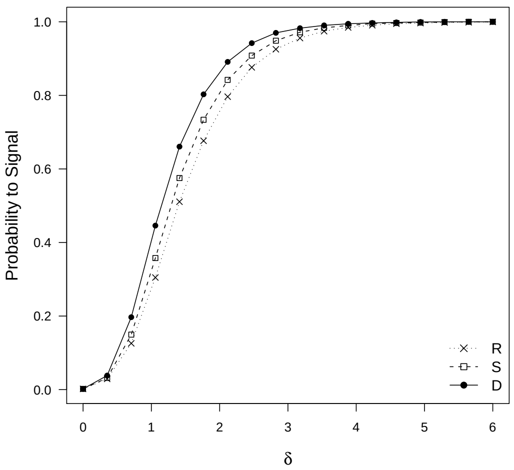
(c) $n = 15$

Figure 3: Power curves of D, R and S charts for $n = 5, 10$ and 15 under Gamma distribution when $\alpha = 0.002$

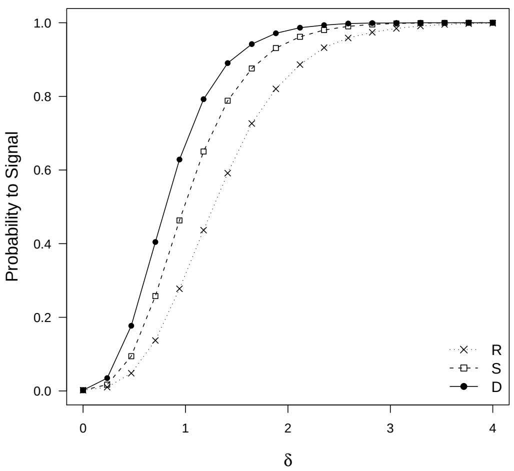
ISBN: 978-988-19251-7-6
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)



(a) $n = 5$



(b) $n = 10$



(c) $n = 15$

Figure 4: Power curves of D, R and S charts for $n = 5, 10$ and 15 under Weibull distribution when $\alpha = 0.002$

ISBN: 978-988-19251-7-6
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)