

2013

Constructing SPC Charts by Cumulative Square and Cumulative Variance Methodologies

Xia Pan

Jeffrey E. Jarrett
University of Rhode Island, jejarrett133@outlook.com

Follow this and additional works at: https://digitalcommons.uri.edu/cba_facpubs

Citation/Publisher Attribution

Pan, X., & Jarrett, J. E. (2013). Constructing SPC Charts by Cumulative Square and Cumulative Variance Methodologies. *International Journal of Economics and Management Engineering*, 3(6), 247-253.

This Article is brought to you by the University of Rhode Island. It has been accepted for inclusion in College of Business Faculty Publications by an authorized administrator of DigitalCommons@URI. For more information, please contact digitalcommons-group@uri.edu. For permission to reuse copyrighted content, contact the author directly.

Constructing SPC Charts by Cumulative Square and Cumulative Variance Methodologies

The University of Rhode Island Faculty have made this article openly available.
Please let us know how Open Access to this research benefits you.

This is a pre-publication author manuscript of the final, published article.

Terms of Use

This article is made available under the terms and conditions applicable towards Open Access Policy Articles, as set forth in our [Terms of Use](#).

**Constructing SPC Charts by Cumulative Square and Cumulative
Variance Methodologies**
(April 2013)

Xia Pan, Ph.D.
Lingnan College, Sun Yat-Sen University

Guangzhou, China
panpapers@yahoo.com

Jeffrey E. Jarrett, Ph.D.*
Management Science, CBA
University of Rhode Island
Kingston, RI 02881, USA
jejarret@mail.uri.edu *

***Corresponding Author**

Abstract

The authors discuss the methodologies for constructing quality control charts by cumulative square and cumulative variance methodologies. These methodologies correspond to the EWMS and EWMV methods proposed by MacGregor and Harris (1993). Based on the distributions of the quality control chart statistics, we focus on the average (mean) cumulative square and average (mean) cumulative variance schemes to arrive at control limits appearing as parallel lines seen in stand Shewhart control charts. Last , we utilize the Box quadratic form and approximation to deal with the complexities of cumulative variance where means vary over time.

Keywords: SPC, cumulative square, cumulative variance, EWMS, EWMV.

Introduction

In statistical process control (SPC), we identify methods based on cumulative statistics as cumulative sum (CUSUM). Previously, many others studied CUSUM control charts including Page (1954, 1961), Johnson (1962a, 1962b), Lucas (1976, 1982, 1985), Hawkins (1981, 1987, 1993), Woodall and Ncube (1985), Crosier (1988), Woodall and Adams (1993), Hawkins and Olwell (1997), Huesch *et al.* (2008, where human generated data indicate that standard Shewhart control charts are not robust) and Ryu, JH *et al.* (2010). Details as to the process for solving quality problems with CUSUM methods detailed by Montgomery (2013, chapter 9) provide for the solution of difficult situations where standard mean and variation control may not find small changes in processes.

The basic idea of CUSUM is to cumulate the information from previous sample points to reflect the changes in the process parameters achieving greater sensitivity. This approach is particularly effective for the sample size $n=1$, where the statistic is a random walk:

$$C_i = \sum_{j=1}^i (x_j - \mu_0) = x_i - \mu_0 + C_{i-1}.$$

The CUSUM statistic is nonstationary. It drifts away from the target as long as one collects a sufficient number of samples, even if there is no shift in the process mean. To reduce the effect of nonstationary, a 'buffer', called reference or slack value, is usually included in practice.

CUSUM chart and EWMA chart, the latter of which studied earlier by Roberts (1959) and Lucas and Saccucci (1990), often have similar records of performance in monitoring statistical processes. MacGregor and Harris (1993) suggested the use exponentially weighted moving variance (EWMV) and exponentially weighted moving square (EWMS) for such tasks. EWMS include squaring values of basic quantities in calculating the weighted moving average. However, previous research tended to focus on the process mean and not the process variability (or variance). Thus, there are no control charts for variability corresponding to MacGregor-Harris's EWMS and EWMV charts. We discuss, as an alternative, cumulative square (CUS) and cumulative variances (CUV) control charts to achieve this goal. The relation of CUS and CUV is similar to the relation of EWMS and EWMV, where the square are cumulated on the basis of grand mean, and the variances are cumulated on the basis of local mean. We further propose another modification of the average cumulative square (ACUS) and average cumulative variances (ACUV) charts.

Cumulative Square and Average Cumulative Square

Cumulative square (CUS) quantity accumulates the variances of the sequential observations of $x_i \sim N(\mu, \sigma_0)$ to form the statistic:

$$C_i^2 = (x_i - \mu)^2 + C_{i-1}^2 \quad i=1, 2, \dots \quad (1)$$

As with the EWMS, μ is the unconditional process mean. Observations x_i arise from independent processes, and each $(x_i - \mu)^2$ is $\sigma_0^2 \chi^2(1)$. Therefore, C_i^2 follows $\sigma_0^2 \chi^2(i)$, as long as C_0^2 is set at zero. The degrees of freedom are the number of observations accumulated. To construct a CUS chart, we determine the control limits through use of the upper and lower critical values of the chi-square distribution with degrees of freedom of i , for example, $\chi_{0.99}^2(i)$ and $\chi_{0.01}^2(i)$.

Figures 1 (a) and (b) show an example of the CUS chart for both variance and mean shift. Since the number of observations accumulated, i , always increases before interruption due to alarm signal, the control limits and the statistic are also monotonously increasing along with i . The chart is lean from the bottom left to the upper right. Moreover, the width between UCL and LCL also increases.

One may prefer control charts with a constant width between the upper and lower control limits for ease of interpretation. With this preference, we utilize the simple mean of the cumulative variance, C_i^2 / i . This is the average chi-square per degree of freedom. We assume the process is in-control in terms of no variance shift, hence, C_i^2 / i follows the distribution $\frac{\sigma_0^2}{i} \chi^2(i)$. However, for large numbers of degrees of freedom i , the chi-square distribution is, according to central limit theorem, approximately a normal

distribution of mean i and variance $2i$. The distribution $\frac{\sigma_0^2}{i} \chi^2(i)$ converges to a constant σ_0^2 and the asymptotic variance becomes zero. Hence, the C_i^2/i control chart is not available for very large i , because the UCL and LCL tend to be the same. In turn, we consider the control chart $C_i^2/\sqrt{i} \sim \frac{\sigma_0^2}{\sqrt{i}} \chi^2(i)$. We may name this chart as average cumulative square (ACUS) for it is the chi-square per square root of degrees of freedom. This statistic approximates to asymptotic distribution of $\sigma_0^2 N(\sqrt{i}, 2)$. It is still a non-horizontal chart, as the slope of the chart is not zero, but the asymptotic width between UCL and LCL is constant. Although the shape of the chart is parabolic curve with the central line \sqrt{i} , it is actually quite flat when i is large, as showed in Figure 1 (c). A further modification may be make ACUS horizontal, that is, to chart $C_i^2/\sqrt{i} - \sigma_0^2 \sqrt{i}$, which follows $\frac{\sigma_0^2}{\sqrt{i}} \chi^2(i) - \sigma_0^2 \sqrt{i}$. This occurs because the mean of $\frac{\sigma_0^2}{\sqrt{i}} \chi^2(i)$ is $\sigma_0^2 \sqrt{i}$. We can still call this ACUS, or horizontal cumulative square (HCUS), shown in Figure 1(d). We estimate the variance of the underlying process, σ_0^2 in Phase I of the SPC control chart construction.

When the number of observations is large, we acquire the control limits through the approximation of chi-square to normal distribution. Fisher's approximation is

$$\chi_\alpha^2 \approx \frac{1}{2}(\sqrt{2i-1} + z_\alpha)^2 \quad (2)$$

for $i > 30$, where z is standard normal distribution. Wilson and Hilferty's (1931) developed a better approximation:

$$\chi^2(i) = i\left(1 - \frac{2}{9i} + \sqrt{\frac{2}{9i}}z\right)^3. \quad (3)$$

This approximation is more accurate than Fisher's approximation for constructing control limits (Hald, 1952).

We base the above discussion on cumulative square on the assumption that we know the process mean μ . When the process mean is unknown, one uses an appropriate estimate to replace μ in the above formulas. In these case, one collects (samples) a set of preliminary data to estimate the process mean. The cumulative square will be composed of sum of $(x_i - \bar{x})^2$ where each of $(x_i - \bar{x})^2 \sim \sigma_0^2 \chi^2(1)$ is not independent of each other. Therefore, the distribution is not exactly $\sigma_0^2 \chi^2(i)$. Moreover, we use approximations (2) and (3), the final approximate error to the control limits may not estimate the process mean well.

Cumulative variances (CUV) and Average Cumulative variances (ACUV)

CUS and ACUS signal both mean shift and variance shift. If one wishes to monitor the shifts in the process variance, one can utilize the concepts of cumulative variances (CUV) or average cumulative variances (ACUV). The cumulative variances (CUV) statistic, corresponding to EWMV in MacGregor and Harris (1993), is defined as

$$C_i^2 = (x_i - \mu_i)^2 + C_{i-1}^2, \quad i=1, 2, \dots \quad (4)$$

where μ_i is the process mean at the i^{th} point. This statistic accumulates the instantaneous variation from the process mean that may vary for different time points. One of the options to estimate μ_i is using EWMA, $\mu_i = (1 - \lambda)\mu_{i-1} + \lambda x_i$ with $\mu_0 = \bar{x}_j$, $j \leq 0$, like in

EWMV. Another option is to measure the unweighted average of moving window of size

$$n, \hat{\mu}_i = \frac{1}{n} \sum_{j=i-n}^i x_j.$$

If μ_i is known, each of $(x_i - \mu_i)^2$ is $\sigma_0^2 \chi^2(1)$, and C_i^2 becomes $\sigma_0^2 \chi^2(i)$. If μ_i is EWMA with parameter λ , this relation may reduce the number of degrees of freedom.

To identify the appropriate number of degrees of freedom of C_i^2 , we apply Box approximation $g\chi^2(v)$ to CUV definition in (4) [This method was employed by MacGregor and Harris (1993) for estimating EWMV].

Let $X' = (x_i, x_{i-1}, \dots, x_1)'$ and $\bar{\mu}' = (\mu_i, \mu_{i-1}, \dots, \mu_1)'$. Since $\mu_i = (1 - \lambda)\mu_{i-1} + \lambda x_i$, set $\mu_0 = 0$, then $\bar{\mu} = LX$, where L is the matrix composed of λ in (5).

$$L = \begin{bmatrix} 0 & \lambda & \lambda(1-\lambda) & \dots & \lambda(1-\lambda)^{i-1} \\ 0 & 0 & \lambda & & \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ & & \ddots & 0 & \lambda \\ 0 & & \dots & 0 & 0 \end{bmatrix} \quad (5)$$

Setting $C_0 = 0$, we have

$$C_i^2 = (x_i - \mu_i)^2 + C_{i-1}^2 = (X' - \bar{\mu}')(X - \bar{\mu}) = X'(I - L)'(I - L)X \quad (6)$$

According to Box (1954) quadratic form, C_i^2 / σ_0^2 follows $g\chi^2(v)$ where $g = \text{trace}(UU) / \text{trace}(U)$ and $v = \text{trace}(U)^2 / \text{trace}(UU)$ and $U = (I - L)'(I - L)$. The values for g and v for different i , the number of accumulated observations, and different values of EWMA parameter λ are shown in Table 1 and Figure 2. From Table 1 and Figure 2, we can see that the value of g and the slope of v/i are almost a constant for different i when i is greater than 10, although the degrees of freedom are less than i .

In addition, we observe that the effects of λ on g and v are almost linear. For small λ , i.e. $\lambda=0.1$, the loss in degrees of freedom is very small, only less 3%. For λ as large as 0.9, the loss in degrees of freedom is as high as 30%.

We analytically calculate the converging asymptotic value of g and v/i because we can express $\text{trace}(U)$ and $\text{trace}(UU)$ with λ and i . In Table 1, the values of g and v/i are constant up to two decimal places. This provides reasonably approximations to g and v/i . For example, when $\lambda=0.1$, approximating $v=0.973i$ and $g=0.873$ are sufficient. This allows us to acquire good approximations for g and v . In turn, the nonstandard chi-square distribution will result in an approximate chi-square.

The choice of λ depends on the purpose of the chart. If the CUV is monitoring variation at the process time-varying mean, the nature of the time-varying mean of the process is less volatile than the process observation. Moreover, λ is small. If the chart is to monitor the variance of forecasting error, then λ is a value between zero and one.

In turn, the CUV statistic C_i^2 follows $\sigma_0^2 g \chi^2(v)$ where the value of g and v depend only on i and λ . For approximations g and v are simply asymptotic values. For example, if $\lambda=0.1$, to ascertain the values of the chi-square, we follow Wilson and Hilferty's (1931):

$$\chi_\alpha^2(v) = v \left(1 - \frac{2}{9v} + \sqrt{\frac{2}{9v}} z \right)^3 \quad (7)$$

The method in (7) is similar to Johnson (1949) in estimating EWMV.

Since v increases with i at a constant rate, the central line and control limits for the control chart also increase. The distance between UCL and LCL widens at a rate

approximately with the square root of i . We observe this in (7) or by a Fisher approximation. For example, if ν is sufficiently large, by Fisher approximation, we have

$$\chi_{0.99}^2(\nu) - \chi_{0.01}^2(\nu) \approx \frac{1}{2}[z_{0.99}^2 - z_{0.01}^2] + 2(z_{0.99} - z_{0.01})\sqrt{2\nu - 1} \quad (8)$$

Therefore, we construct a control chart with parallel lines for control limits and of constant width for the conditional variance that is similar to the ACUS. Finally, the

ACUV chart statistic is $C_i^2 / \sqrt{i} \sim \frac{\sigma_0^2}{\sqrt{i}} g \chi^2(\nu)$, because $\nu = (\frac{\nu}{i})i$ and $\frac{\nu}{i}$ is nearly constant.

Figure 3 shows examples of the CUV and ACUV chart for $\lambda=0.1$ choosing values of $g=0.873$ and $\frac{\nu}{i}=0.97$. We construct control limits at a two percent significance level ($\alpha=0.02$). A mean shift from 0 to 1σ on Obs. #31 and a variance shift from 1σ to 2σ on obs.#45 occurred in the process. We observe that the CUV charts do not reflect the shift in mean but perform well in detecting the shift in variance. The ACUV control chart performed well. Therefore, these control charts are very useful for monitoring variability of processes.

Summary

We observed that the CUS, ACUS, CUV, and ACUV are easier to construct in comparison with EWMS and EWMV control charts. Another advantage of CUS and ACUS in comparison with EWMS is that $\chi^2(i)$ is exact for a process that is in control. When the chi-square distribution is exact for a process that is in control, one can easily construct control limits from public available distribution tables and computer functions.

The CUV and ACUV also have advantage in that the approximation procedure can employ the nearly constant characteristics of g and $\frac{v}{i}$ in order to construct control limits easily. [One may easily accomplish this task with spreadsheet software, but one must beware on computational errors common to spreadsheet software.] The ACUV control chart contains a constant distance between the UCL and LCL and appears similar to conventional Shewhart control charts.

In addition, we illustrated our results rather than only producing tables with numerical values. This eases our understandings that control limits are not constant over time when sample (subgroup) sizes are the same. In the future, we shall consider the relative performance of the scheme by assessing the criterion of average run length (ARL). As the cumulative square or cumulative variances are serially correlated, the ARL of these schemes does not directly reflect the significance level α associated with the control limits. We expect to obtain ARL through Monte Carlo simulation for processes that are in varying degrees of control and for various changes in process variability.

References

- Box, G. E. P, (1954) Some Theorems on Quadratic forms Applied in the Study of Analysis of Variance Problems: Effect of Inequality of Variance in One-Way Classification. *Annals of Mathematical Statistical Statistics*, 25, 290-302.
- Crosier, R.B. (1988) "Multivariate Generalizations of Communicative Sum Quality Control Schemes," *Technometrics*, 30 (3), 291-303.
- Hald, A. (1952) *Statistical Theory with Engineering Applications*, John Wiley & Sons, Inc. New York
- Hawkins, D.M. and D.H. Olwell (1997) *Cumulative Sum Charts and Charting for Quality Improvement*, New York: Springer-Verlag
- Hawkins, D.M.(1981) "A CUSUM for a Scale Parameter", *Journal of Quality Technology*, 13 (4), 228-235.
- Hawkins, D.M.(1987) "Self-Starting CUSUM Charts for Location and Scale", *The Statistician*, 36 (4), 299-316.
- Hawkins, D.M.(1993) "Cumulative Sum Control Charting: An Underutilized SQC Tool", *Quality Engineering*, 5 (3), 463-477.
- Huesch, M.D., Madan, A. and J.J. Borkhardt (2008) 'An empirical evaluation of cusum control charts for use in robust quality control of surgical and binary outcomes,' *Quality Managment in Health Care*, 17, (3), 218-226.
- Johnson, N. L. (1949) "System frequency Curves Generated by Methods of Transflation," *Biometrika* 36, 149.
- Johnson, N. L. and F.C. Leone (1962a) " Cumulative Sum Control Charts--Mathematical Principles Applied to Their Construction and Use", Part I, *Industrial Quality Control*, 19 (2), 15-21.
- Johnson, N. L. and F.C. Leone (1962b) " Cumulative Sum Control Charts--Mathematical Principles Applied to Their Construction and Use" Part II, *Industrial Quality Control*, 19 (19), 22-28.
- Lucas, J.M. (1976) "The design and use of quality control schemes," *Journal of Quality Technology*, 8 (1), 1-12.
- Lucas, J.M., and M. S. Saccucci (1990) "Exponentially weighted Moving Average Control Schemes: Properties and Enhancements", *Technometrics*, 32 (1), 1-29.
- MacGregor, J. F. and T.J. Harris (1993) "The Exponentially Weighted Moving Variance", *Journal of Quality Technology*, 25 (2), 106-118.

Montgomery, D.C. (2013) *Introduction to Statistical Quality Control* ^{7th ed.}, John Wiley & Sons.

Roberts, S.W. (1959) "Control Chart Tests Based on Geometric Moving Averages," *Technometrics*, 42 (1), 97-102.

Ryu, JH, Wan, H. and S. Kim (2010) "Optimal Design of a CUSUM Control Chart for Mean Shift of Unknown Size," *Journal of Quality Technology*, 42, (3), 1-16.

Wilson, E.B. and M. M. Hilferty (1931) *The Distribution of Chi-Square*, Proc. Nat.Acab.Sci. Washington 17, 684-688.

Woodall, W. H. and B. M. Adams (1993) "The Statistical designs of CUSUM Charts", *Quality Engineering*, 5 (4), 559-570.

Woodall H.W. and M.M. Ncube (1985) "Multivariate CUSUM Quality Control Procedures", *Technometrics*, 27(3), 285-292.

Table 1

λ	g		v		Ratio of v/i		
	i =30	i =100	i =30	i =100	i =30	i =100	i =300
0.1	0.8728	0.8744	29.0503	97.2519	0.9683	0.9725	0.9737
0.5	0.387	0.3883	25.5516	85.5512	0.8517	0.8555	0.8566
0.9	0.0254	0.0255	21.1673	70.8434	0.7056	0.7084	0.7093

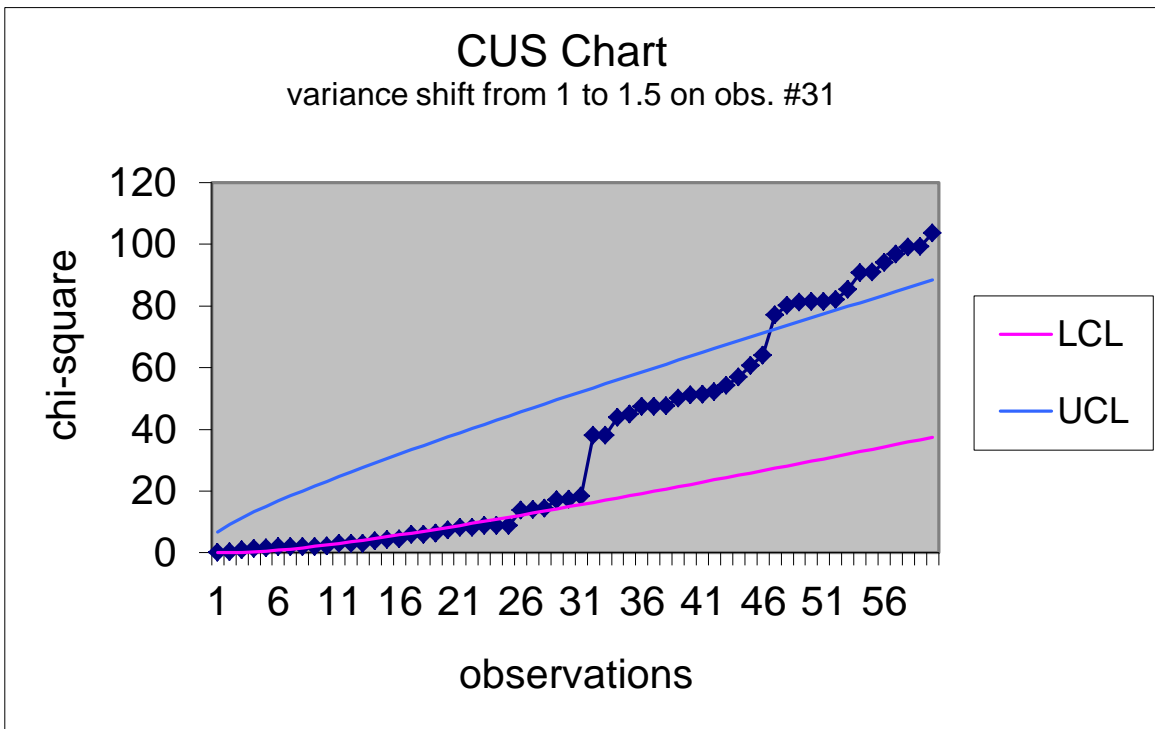


Figure 1 (a) CUS Chart with variance shift

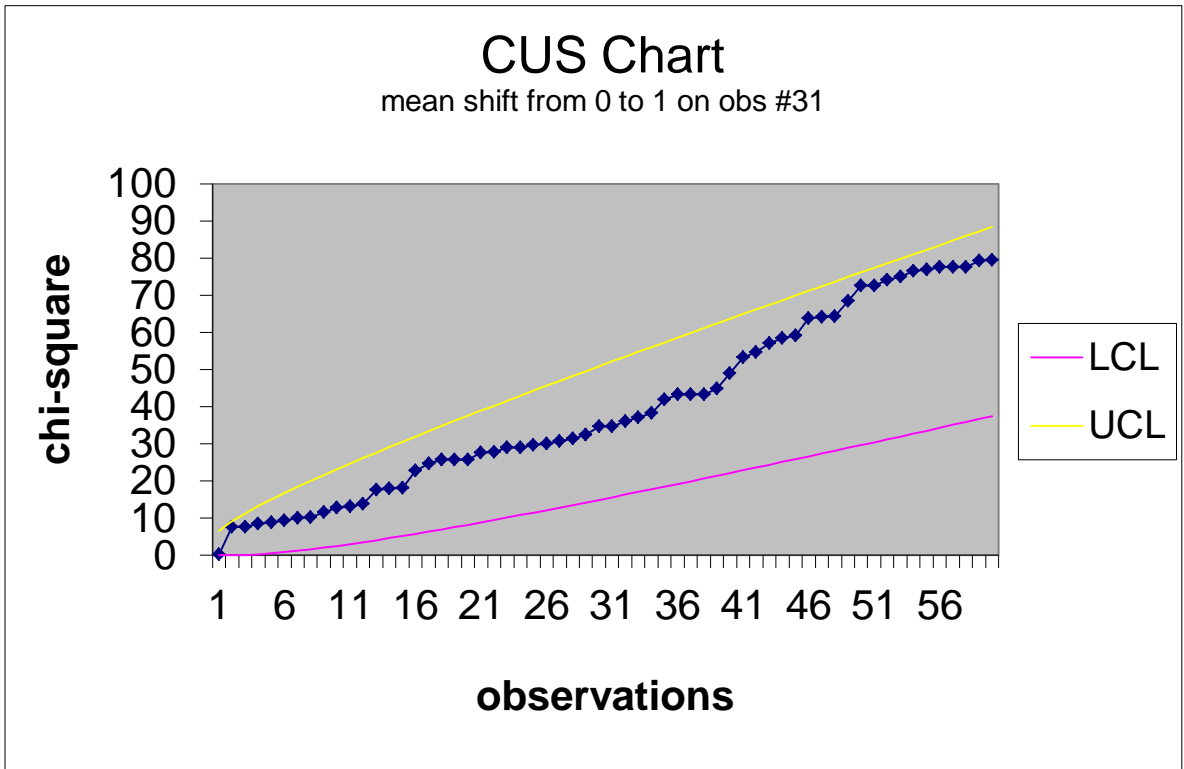


Figure 1 (b) CUS Chart with mean shift

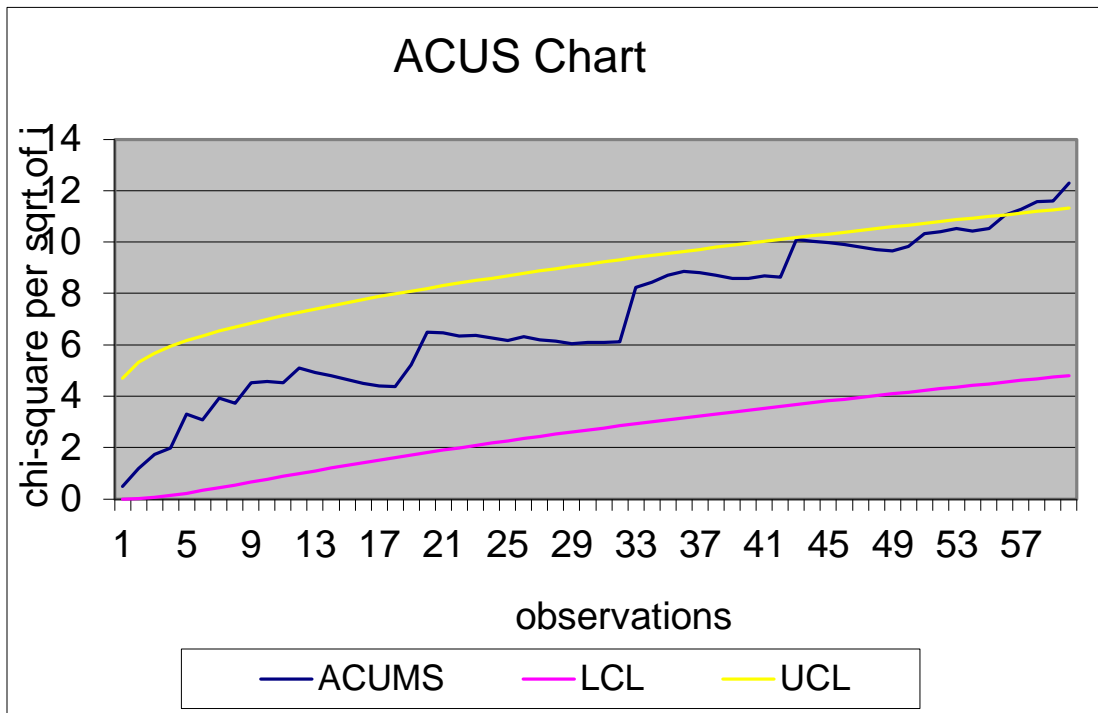


Figure 1 (c) ACUS Chart

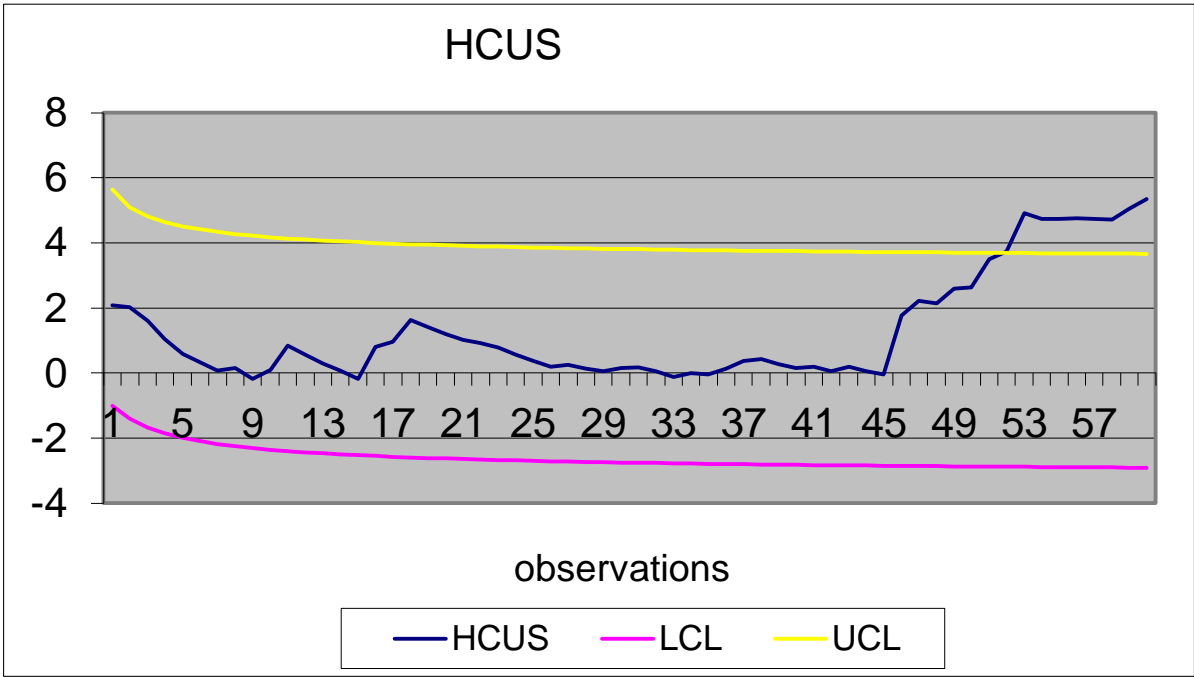


Figure 1 (d) HCUS Chart with mean shift at Obs. # 31

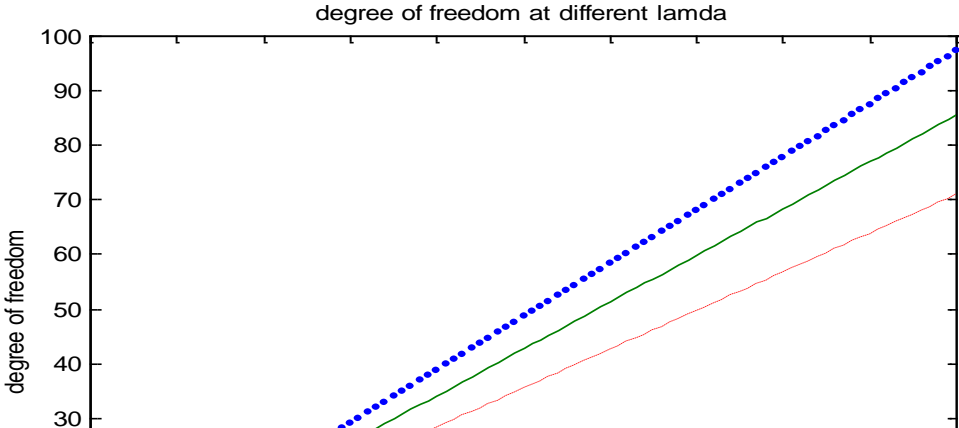


Figure 2 (a) degree of freedom of C_i^2 / σ_0^2

dots: $\lambda = 0.1$; line: $\lambda = 0.5$; dash: $\lambda = 0.9$

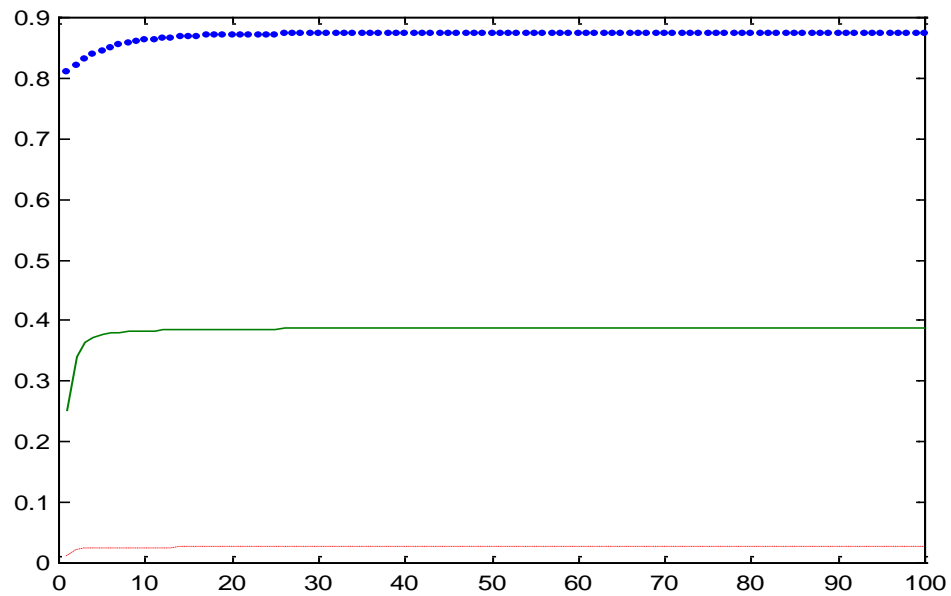


Figure 2 (b) value of g

dots: $\lambda = 0.1$; line: $\lambda = 0.5$; dash: $\lambda = 0.9$

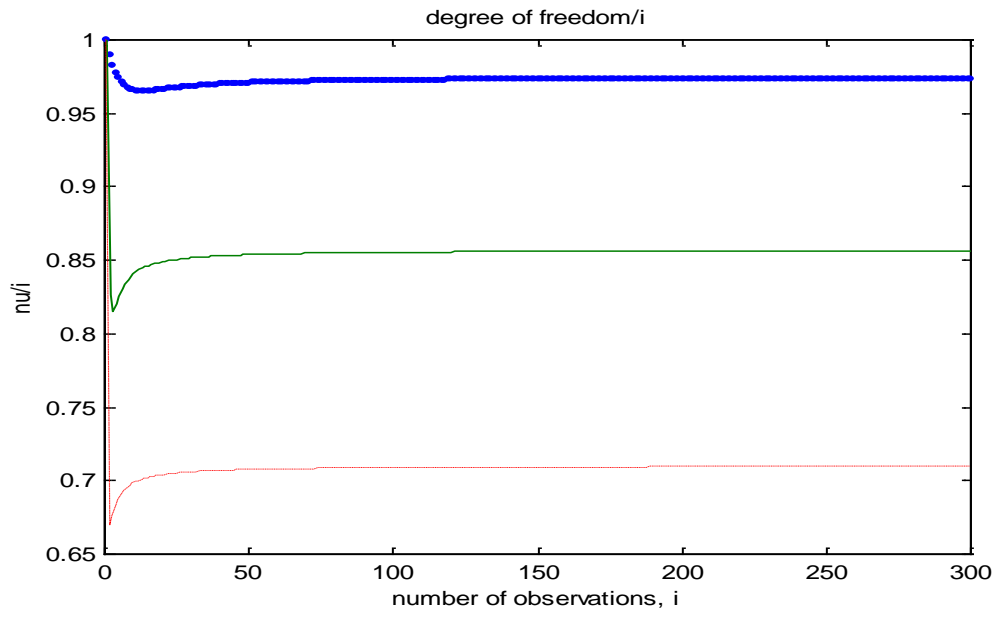


Figure 2 (c) ratio of ν/i

dots: $\lambda = 0.1$; line: $\lambda = 0.5$; dash: $\lambda = 0.9$

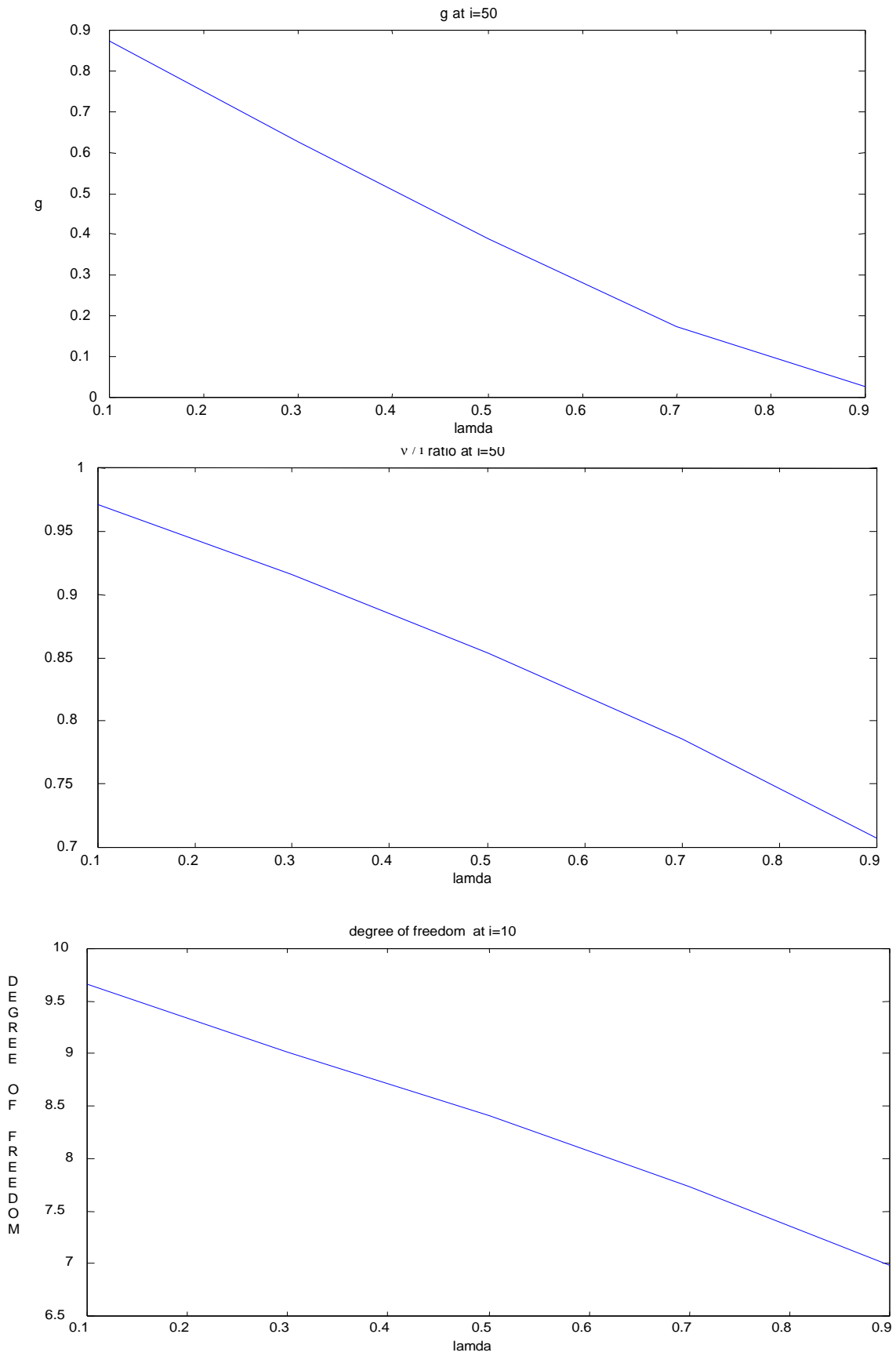


Figure 2 (d) The Effects of λ

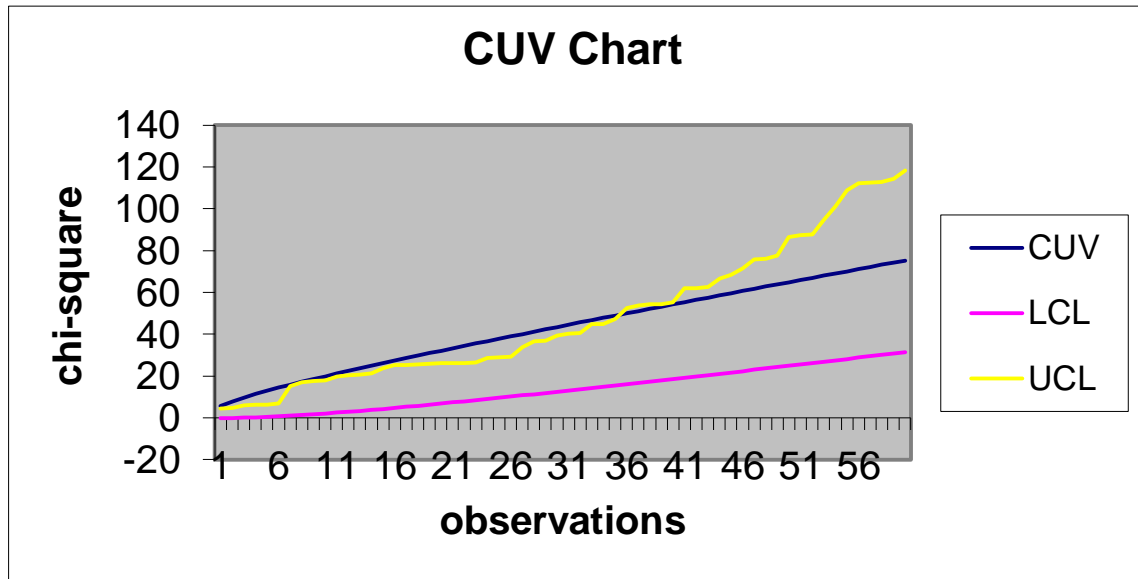


Figure 3 (a) ACUV Chart with $\lambda = 0.1$

Mean shift from 0 to 1σ on Obs. #31, variance shift from 1σ to 2σ on obs.#45

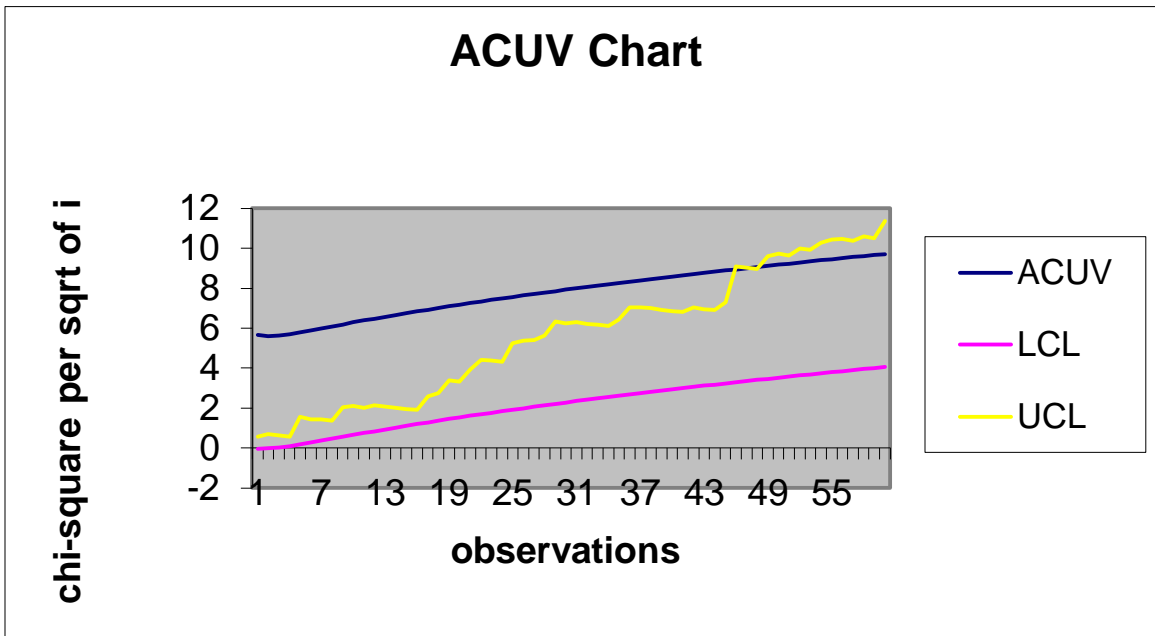


Figure 3 (b) ACUV Chart with $\lambda = 0.1$

Mean shift from 0 to 1σ on Obs. #31, variance shift from 1σ to 2σ on obs.#45