Tamsui Oxford Journal of Information and Mathematical Sciences **27(2)** (2011) 117-136
Aletheia University

An Investigation of the Effects of Misclassification Errors on the Analysis of Means *

S. Balamurali†

Department of Mathematics,Kalasalingam University
Anand Nagar, Krishnankoil-626126, India

and

M. Kalyanasundaram

STMicroelectronics Pte Ltd, #18, Level - 2,
Ang Mo Kio Industrial Park -2, Singapore - 569-505

Received November 17, 2005, Accepted December 22, 2010.

Abstract

In this paper, mathematical and graphical evaluation of the effects of misclassification error on the Analysis of Means is given. The operating characteristic curve under different error rates is explained and illustrated.

Keywords and Phrases: *Analysis of means, Attribute data, Misclassification errors, Multiple comparison, Variables data.*

---

*2000 *Mathematics Subject Classification.* Primary 62D05, 62J10, 62P30.

†E-mail: sbmurali@rediffmail.com

1. Introduction

The methodology of Analysis of Means (ANOM) was developed and introduced into the literature by Ott (1967). Later, Schilling (1973a, b) extended the ANOM concept to address interaction and main effects for a variety of experimental designs. Schilling called his extension as the Analysis of Means for Treatment Effects (ANOME), where k means being compared are not necessarily independent. Nelson (1983a) gives necessary mathematical constants. Balamurali (2010) provided simplified factors of ANOME constants. The ANOM is sometimes referred to as an alternative to the Analysis of Variance (ANOVA). Its advantage over the ANOVA procedure is that the data are plotted and thus the results may be quickly interpreted. The ANOM technique provides a control chart-like approach to the analysis of experimental data. Consequently, it is to be expected that the increasing interest in control chart should lead to an increased interest in ANOM. Control charts under error are currently generating much interest and a growing body of literature. The purpose of this paper is, therefore, to determine and illustrate the effect of misclassification error on the operating characteristic (OC) curve of ANOM.

2. Attributes Data

Ott (1975) had pointed out that important problems arise in almost every industrial process where the economically critical characteristics of the product are attributes. The ANOM was extended to the case where the normal distribution can be used as an approximation to the actual distribution of the data or one can transform the data to make it approximately normal. Transformation of attribute data is discussed in Nelson (1983b).

2.1. Proportions Data

When the data consists of the number or proportion of units having a particular attribute, it can often be represented by a binomial distribution. The procedural steps are outlined below.

(i) Obtain samples of equal size n from each of k populations. Let the number of units having the attribute of interest in each of the k samples be denoted by $X_1, X_2, \dots, X_k$.

(ii) Compute the k sample proportions $p_i = \frac{X_i}{n}, i = i, 2, \ldots, k$.

(iii) Compute the overall sample proportion $\overline{p} = \sum_{i=1}^{k} \frac{p_i}{k}$.

(iv) Compute an estimate of the standard deviation of the sample proportion using $S_{\overline{p}} = \sqrt{\frac{\overline{p}(1-\overline{p})}{n}}$.

(v) Determine the decision lines at risk α as

$$\begin{aligned} \text{Upper Decision Line(UDL)} &= \overline{p} + h_\alpha s_{\overline{p}}\sqrt{\frac{(k-1)}{k}} \\ \text{Center Line (CL)} &= \overline{p} \\ \text{Lower Decision Line (LDL)} &= \overline{p} - h_\alpha s_{\overline{p}}\sqrt{\frac{(k-1)}{k}} \end{aligned} \tag{1}$$

where h_α is the table value in Nelson (1983a) for specified value of α and k means with infinite degrees of freedom. The infinite degrees of freedom used because one is approximating the binomial distribution with a normal distribution.

(vi) Plot the sample proportions p_i against the decision lines. If all proportions fall between the decision lines, then accept the hypothesis of k equal proportions. Otherwise conclude that the excessive variability exists and the process is not in statistical control. When the standard or target value is given by p, the decision lines are obtained by

$$\begin{aligned} \text{Upper Decision Line (UDL)} &= p + h_\alpha\sqrt{\frac{p(1-p)}{n}} \\ \text{Center Line (CL)} &= p \\ \text{Lower Decision Line (LDL)} &= p - h_\alpha\sqrt{\frac{p(1-p)}{n}} \end{aligned} \tag{2}$$

where, as before, n is the number of items inspected and h_α is the value obtained from Nelson (1983a) for specified value of α and k with ∞ degrees of freedom.

When a sample of size n is taken, the sample fraction defective given by the statistic x/n is compared to the decision lines. The largest and smallest values of x for which an in-control state is indicated are;

$$x_l = [nUDL]^- \text{ and } x_s = [nLDL]^+ \tag{3}$$

where $[\ ]^-$ and $[\ ]^+$ indicate round-down and round-up operations respectively, to the nearest integer. The OC curve illustrates the probability that a sample fraction defective x/n will fall within control limits as a function of the true process fraction defective p, i.e.,

$$P_a(p) = \sum_{x=max\{0,x_s\}}^{x_l} \binom{n}{x} p^x (1-p)^{n-x},\ 0 \le p \le 1 \tag{4}$$

Example 1. Consider the following example. (Given in Ott (1975), pp.106).

Table 1: Effect of Copper on Corrosion

Level of Copper (ppm)	Containers Examined n	Failures d_i	Fraction failing p_i
5	80	14	0.175
10	80	36	0.450
15	80	47	0.588
Total	240	97	

The overall sample proportion of failures is $\overline{p} = (p_1 + p_2 + p_3)/3$ =0.404=40.4%.

The standard deviation of the proportion is $S_{\overline{p}} = \sqrt{\frac{\overline{p}(1-\overline{p})}{n}}$=5.5% for $\overline{p} = 0.404$ and $n = 80$. For $k = 3$ and $\alpha = 0.01$, referring to the Nelson (1983a) table, we get $h = 2.91$. Then, $CL = 0.404 = 40.4\%$, $UDL = 0.5344 = 53.44\%$ and $LDL = 0.2736 = 27.36\%$. The ANOM chart is shown in Figure 1.

Using equations (3) and (4) we can calculate the probability P_a that a sample fraction defective x/n will fall within control limit as a function of the true process fraction defective, p. That is.,

$$P_a(p) = \sum_{x=max\{0,x_s\}}^{x_l} \binom{n}{x} p^x (1-p)^{n-x},\ 0 \le p \le 1$$

The values of x_l and x_s are $x_l = [80(0.5344)]^- = 42$ and $x_s = [80(0.2736)]^+ = 22$.

Therefore, the OC curve is expressed mathematically as

$$P_a(p) = \sum_{x=22}^{42} \binom{80}{x} p^x \ (1-p)^{80-x}, \ 0 \leq p \leq 1$$

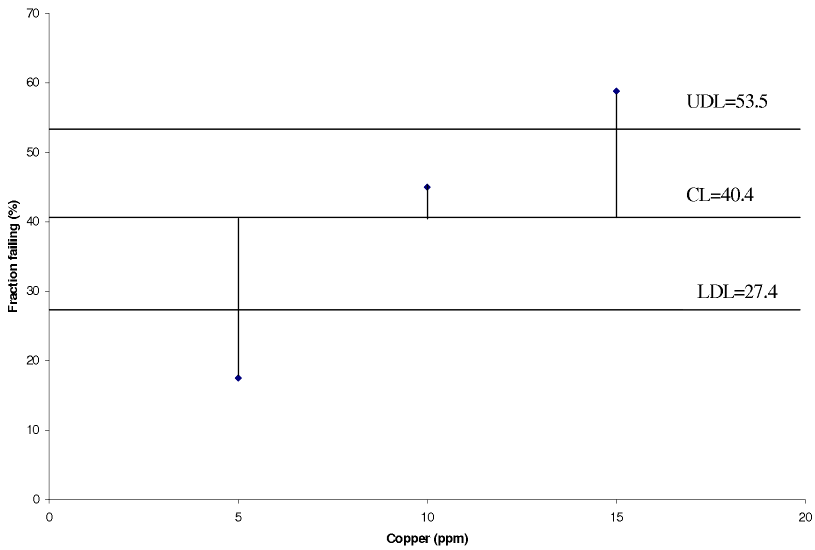

Figure 1: Effect of Copper on Corrosion

2.1.1 Development of Formulae under Misclassification Error

There are only two types of error possible in attribute sampling. An item that is good may be regarded as defective (Type I error) or a defective item may be classified as good (Type II error).

Let $E_1 =$ The event that a good item is classified as a defective item
$E_2 =$ The event that a defective item is classified as good item
$A =$ The event that an item is defective and
$B =$ The event that an item is classified as a defective

Then

$$P(B) \;=\; P(A)\; P(\overline{E}_2) \;+\; P(\overline{A})\; P(E_1) \tag{5}$$

Also by defining the quantities,

$$\begin{aligned}
p \;&=\; P(A) \;=\; \text{true fraction defective}\\
p_e =\; &P(B) \;=\; \text{apparent fraction defective}\\
e_1 =\; &P(E_1) \;=\; \text{the probability that } \mathrm{E}_1 \text{ occurs, and}\\
e_2 =\; &P(E_2) \;=\; \text{the probability that } \mathrm{E}_2 \text{ occurs}
\end{aligned}$$

Then the apparent fraction defective may be expressed as

$$p_e = \; p\;(1-e_2)\; + (1-p)\; e_1 \tag{6}$$

The decision lines under misclassification error when the standards are based upon the data will be modified from equation (1) and the decision lines are given by,

$$\begin{aligned}
\text{Upper Decision Line (UDL)} &= p_e \; + h_\alpha \; s_{p_e}\sqrt{\frac{(k-1)}{k}}\\
\text{Center Line (CL)} &= p_e\\
\text{Lower Decision Line (LDL)} &= p_e \; - h_\alpha \; s_{p_e}\sqrt{\frac{(k-1)}{k}}
\end{aligned} \tag{7}$$

where $S_{p_e} = \sqrt{\frac{p_e(1-p_e)}{n}}$ and $p_e = p(1-e_2) + (1-p)e_1$.

If p is treated as a target value, the center line and decision limits for the ANOM chart will be exactly as given in equation (2). Now the statistic of interest is the sample apparent fraction defective y_e/n, where y_e is the number of apparent defectives observed by the inspector. The largest and smallest values of y_e for which an in-control state is indicated will be

$$y_{el_1} = [nUDL_{P_e}]^- \text{ and } y_{es_1} = [nLDL_{p_e}]^+ \tag{8}$$

Then the OC curve under inspection error is given by

$$P_{ae_1}(p_e) = \sum_{y_e=max(0,y_{es_1})}^{y_{el_1}} \binom{n}{y_e} p_e^{y_e}(1-p_e)^{n-y_e},\ 0 \le p_e \le 1 \tag{9}$$

Clearly, although the ANOM chart remains identical to the error free case, inspection will distort the OC curve of the ANOM chart.

Example 2. Again let us consider Example 1, and assume that the ANOM for proportions is now based upon data observed from an error prone inspection process operating at an actual fraction defective $p = 0.404$ and $\alpha = 0.01$. The four cases of Type I and Type II errors considered are $(0.00, 0.00)$; $(0.05, 0.00)$; $(0.00, 0.05)$ and $(0.05, 0.05)$. The center line and other decision lines will change as a function of each error pair considered. To illustrate the calculations, let us consider $(e_1, e_2) = (0.05, 0.05)$. In this case we have the decision lines as

$$\begin{aligned} \text{Center Line (CL)} &= p_e \ = \ 0.4136 \\ \text{Upper Decision Line (UDL)} &= 0.5444 \\ \text{Lower Decision Line (LDL)} &= 0.2828 \end{aligned}$$

The values of y_{el} and y_{es} are 43 and 23 respectively. The OC curve under this error pair would be given by

$$P_{ae_1}(p_e) = \sum_{y_e=23}^{43} \binom{80}{y_e} p_e^{y_e}(1-p_e)^{80-y_e},\ 0 \le p_e \le 1$$

Table 2 shows the decision line values under four cases of error pairs.

Table 2: Decision Line Values under Error

(e_1, e_2)	CL_{pe}	UDL_{pe}	LDL_{pe}
0.00, 0.00	0.4040	0.5344	0.2736
0.05, 0.00	0.4338	0.5654	0.3021
0.00, 0.05	0.3838	0.5130	0.2546
0.05, 0.05	0.4136	0.5444	0.2828

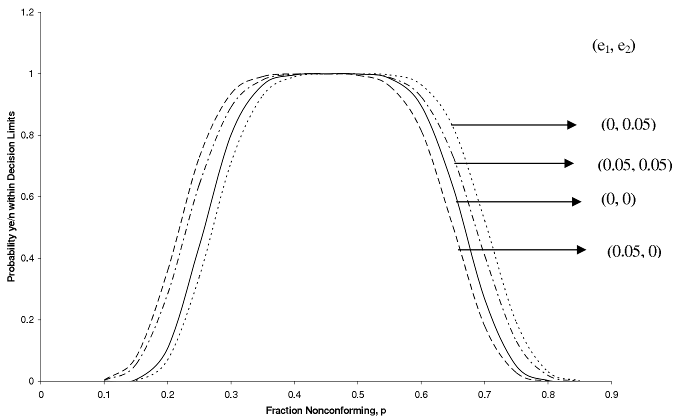

Figure 2: Operating Characteristic (OC) Curves under Different Error Rates

Figure 2 illustrates the OC curve under each error pair considered. It is easily seen that realistic quantities of error have considerable effect on the OC curve. From the graph it is easily observed that Type I error increases the value of P_{ae1} for low process fraction defective and decreases the value of P_{ae1} at high process fraction defective. We already defined Type I error is the erroneous classification of a good item as defective. Therefore, the probability that the sample statistic y_e/n falling above LDL is increased when errors are made. Similarly, Type I error causes the sample statistic to fall above LDL. It is also obvious that Type II error decreases the value of P_{ae1} for low fraction defective and increases the value of P_{ae1} at high fraction defective. When both errors are operative together, the Type I error has more influence on the OC curve for low value of p, while the Type II error dominates the effect on the OC curve for higher values of p. The reason is that when the actual process fraction defective is quite low there is a little chance for a Type II error. When p increases the effect of this error is also increased.

2.1.2 Adjusted OC Curve

The decision lines may be determined for the proportions data, to large extent, compensate for misclassification error and provide an OC curve closer to the desired. First the center line and other decision lines are determined using equations (2). This is easily done based upon the target value or actual fraction defective p. However, the value of p to be used in (1) may be found by solving equation (6) for p.

$$p = \frac{p_e - e_1}{1 - e_1 - e_2} \tag{10}$$

In order to estimate e_1 and e_2, p must be known. This is true, and p may be determined by inspection job samples, post inspection audits etc. However, once floor production sampling begins, p is not known and must be estimated by the apparent process fraction defective p_e.

Once the center line CL_p and other decision lines UDL_p and LDL_p have been determined, the compensating adjusted center line (ACL) and other adjusted decision lines (AUDL and ALDL) are calculated as follows:

$$\begin{aligned} ACL &= CL_p(1-e_2) + (1-CL_p)e_1 \\ AUDL &= UDL_p(1-e_2) + (1-UDL_p)e_1 \\ ALDL &= LDL_p(1-e_2) + (1-LDL_p)e_1 \end{aligned} \tag{11}$$

It is to be noted that the compensating ANOM chart simply reflects taking the error free center line and control limit fraction defective, and using equation (6) to convert them to error-prone equivalents.

The largest and the smallest values of y_e for which an in-control state is indicated are

$$y_{ela} = [nAUDL_{pe}]^{-} \text{and } y_{esa} = [nALDL_{pe}]^{+} \tag{12}$$

The adjusted OC curve ($P_{aae}(p_e)$) under misclassification error is given by,

$$P_{aae}(p_e) = \sum_{y_e = Max(0, y_{es_a})}^{y_{el_a}} \binom{n}{y_e} p_e^{y_e}(1-p_e)^{n-y_e}, 0 \le p_e \le 1 \tag{13}$$

2.2 Count Data

ANOM can also be applied to count data (non-conformities) in which the Poisson distribution is an appropriate model. The decision lines are obtained by,

$$\begin{aligned} \text{Upper Decision Line(UDL)} &= \bar{c} + h_\alpha \sqrt{\bar{c}} \sqrt{\frac{(k-1)}{k}} \\ \text{Center Line(CL)} &= \bar{c} \\ \text{Lower Decision Line(LDL)} &= \bar{c} - h_\alpha \sqrt{\bar{c}} \sqrt{\frac{(k-1)}{k}} \end{aligned} \tag{14}$$

where $\overline{c}$ is the overall average and is expressed mathematically as $\overline{c} = \sum_{i=1}^{k} \frac{c_i}{k}$. Let X be the number of non-conformities (defects) in the sample. Then the largest and the smallest values of X for which an in-control state is indicated are

$$x_l = [UDL]^- \quad \text{and} \quad x_s = [LDL]^+ \tag{15}$$

The OC function is given by

$$P_a(\overline{c}) = \sum_{x=Max(0,x_s)}^{x_l} \frac{e^{-\overline{c}}\overline{c}^x}{x!}, \overline{c} > 0 \tag{16}$$

where $P_a(\overline{c})$ represents the probability that a sample point (number of non-conformities) will fall within the decision limits. If the standard value (target value) is available by c, then the decision lines are given by

$$\begin{aligned} \text{CL} &= c \\ \text{UDL} &= c + h_\alpha\sqrt{c} \\ \text{LDL} &= c - h_\alpha\sqrt{c} \end{aligned} \tag{17}$$

Again the largest and the smallest values of X, for which an in-control state is indicated are $x_l = [UDL]^-$ and $x_s = [LDL]^+$. The OC function is then given by

$$P_a(c) = \sum_{x=Max(0,x_s)}^{x_l} \frac{e^{-c}c^x}{x!}, \ c > 0$$

Example 3. Consider the following problem (given in Wadsworth et al. (1986) pp.614).

Sample Number	Number of Non-conformities
1	11
2	23
3	35
4	19
5	22
6	25
7	28
8	14
9	50
10	23
	250

The average count $\overline{c} = \frac{250}{10} = 25$; $s = \sqrt{25}$; h_α for $\alpha = 0.05, k = 10$ and for ∞ degrees of freedom is 2.8. The decision lines computed using equation (14) are, $25 \pm 2.8(5)\sqrt{9/10}$. Thus $UDL_c = 25 + 13.3 = 38.3, CL_c = 25$ and $LDL_c = 25 - 13.3 = 11.7$.

2.2.1 Development of Formulae under Misclassification Error

In using ANOM for count data, a sample is taken and the number of non-conformities is counted. Again the two types of error are (i) failing to find one or more number of non-conformities in the sample and (ii) declaring one or more non-conformities when none exist (a false alarm). Let

$$\begin{aligned} u &= \text{probability that a non}-\text{conformity is correctly noted} \\ v &= \text{the average number of false alarm per part} \\ c &= \text{true average number of non}-\text{conformities per part} \\ c_0 &= \text{average number of defects per part observed by the inspector.} \end{aligned}$$

Then we have

$$c_0 = uc + v \tag{18}$$

with both u and v estimated. Every effort should be made to eliminate both types of errors, i.e., to get u close to one and v close to zero. Thus we obtained equation (17) as that of Ran Suich (1988). If the ANOM for count is based upon past data which is subject to misclassification error, then the equations

(14) and (17) can be rewritten on the basis of estimated value of c_0 instead of c. Therefore, the error prone and error free decision lines will be

$$\begin{aligned} \mathrm{CL}_{c_0} &= c_0 \\ \mathrm{UDL}_{c_0} &= c_0 + h_\alpha \sqrt{c_0}\sqrt{\frac{(k-1)}{k}} \\ \mathrm{LDL}_{c_0} &= c_0 - h_\alpha \sqrt{c_0}\sqrt{\frac{(k-1)}{k}} \end{aligned} \tag{19}$$

and

$$\begin{aligned} CL_0 &= c \\ UDL_0 &= c + h_\alpha \sqrt{c} \\ LDL_0 &= c - h_\alpha \sqrt{c} \end{aligned} \tag{20}$$

respectively. The OC function under error prone data is given by

$$P_a(c_0) = \sum_{x=Max(0,x_s)}^{x_l} \frac{e^{-c_0} c_0^x}{x!} \tag{21}$$

where $x_l = [UDL_{c_0}]^-$ and $x_s = [LDL_{c_0}]^+$.

Example 4. Consider the example 3, under misclassification error. Here we consider four cases $(u, v) = (1, 0), (1, 2), (0.8, 0)$ and $(0.8, 2)$. The first case corresponds to sampling without misclassification error while the other three represent different error rates. The following table shows different decision limits under different error rates and the graph shows the effects on OC curve.

Table 3: Decision Lines under Misclassification Error

u	v	UDL_c	CL_c	LDL_c
1.0	0.0	38.3	25.0	11.7
1.0	2.0	40.803	27.0	13.2
0.8	0.0	31.88	20.0	8.12
0.8	2.0	34.46	22.0	9.54

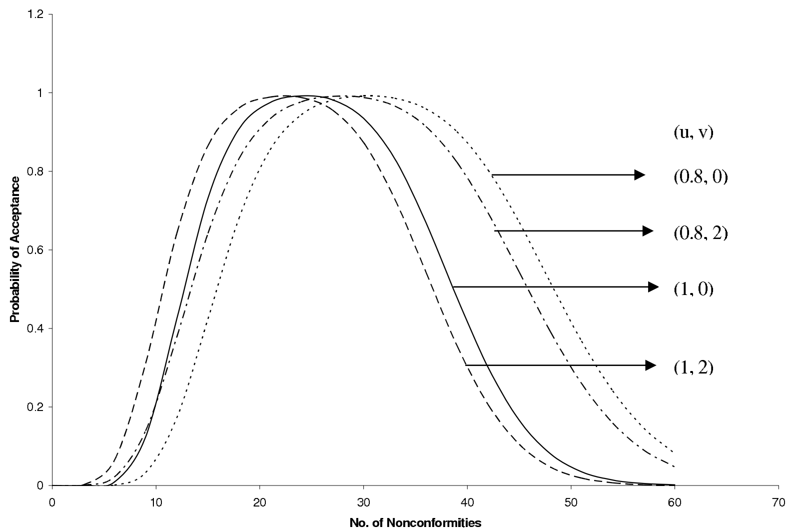

Figure 3: Operating Characteristic (OC) Curves under Different Error Rates

From Figure 3, it can be seen that (1, 2) results in a shifting of the OC curve to the left. This is because $u = 1$ means that none of non-conformities were missed while $v = 2$ means that there will be, on the average, two false alarms per part and this will shift the OC curve to the left side of the true OC curve. An error rate of (0.8, 0) results in huge shift of the OC curve to the right. That is, $u = 0.8$ means that the inspector is finding only 80% of the non-conformities.

2.2.2 Adjusted OC Curve

First the center line and other decision lines are determined using equation (17). This is easily done based upon the target value or actual fraction non-conformities c. However, the value of c to use in the equation (17) may be found by solving the equation (18) for c as

$$c = \frac{\overline{c}_0 - v}{u} \tag{22}$$

We then use the value of c from the equation (22) in equation (17). In either case we obtain the Adjusted Center Line (ACL) and Adjusted Decision Lines (AUDL and ALDL) yes

$$\begin{aligned} ACL_c &= CL_c(u) + v \\ AUDL_c &= UDL_c(u) + v \\ ALDL_c &= LDL_c(u) + v \end{aligned} \tag{23}$$

Then x_{al} and x_{as}, the largest and the smallest values of X, for which the process is in control are

$$x_{al} = [AUDL_c]^{-} \text{and} \quad x_{as} = [ALDL_c]^{+} \tag{24}$$

The adjusted OC curve is given by

$$P_a(c) = \sum_{w=Max(0,x_{as})}^{x_{al}} \frac{e^{-c}c^w}{w!} \tag{25}$$

2.3 Measurement Data

Ramig (1983) describes the ANOM procedure very clearly. We can use her step by step procedure to explain the technique here. As a first example, let us assume that we have samples all of size n. All the samples are assumed to have come from normal population with the same variance σ^2. Let x_{ij} be the j^{th} observation from the population i. Let $\overline{\overline{X}}$ represents the grand mean and s^2 is the pooled estimate of the common but unknown variance. These quantities are defined mathematically by,

$$\overline{\overline{X}} = \sum_{i=1}^{k} \frac{\overline{X}_i}{k} \tag{26}$$

$$s^2 = \sum_{i=1}^{k} \frac{s_i^2}{k} \tag{27}$$

$$\text{where } s_i^2 = \sum_{j=1}^{n} \frac{(x_{ij} - \overline{X}_i)^2}{(n-1)} \text{ and } \overline{X}_i = \sum_{j=1}^{n} \frac{x_{ij}}{n}.$$

The procedure is as given below.

1. Compute the group mean $\overline{X}_i, (i = 1, 2, 3, \ldots, k)$
2. Compute the grand mean $\overline{\overline{X}}$ using equation (26)
3. Compute s, an estimate of the standard deviation of an individual observation. This is the square root of s^2, where s^2 is computed using (27).
4. Obtain the value of h_α , from the table in Nelson (1983a) for the level of significance α , number of means k and degrees of freedom $(n-1)k$.
5. Compute the decision lines as follows

$$\begin{aligned} CL &= \overline{\overline{X}} \\ UDL &= \overline{\overline{X}} + h_\alpha s \sqrt{\frac{(k-1)}{kn}} \\ LDL &= \overline{\overline{X}} - h_\alpha s \sqrt{\frac{(k-1)}{kn}} \end{aligned} \tag{28}$$

6. Plot the means against the decision lines. If any mean falls outside the decision lines then conclude that there is a statistically significant difference among the means.

The decision lines in (28) were obtained under the assumption that μ and σ are unknown, so they were estimated by $\overline{\overline{X}}$ and s respectively, where s is as defined in (27), ν is the degrees of freedom upon which s is based. If μ and/or σ were both known, then they can be used for calculation of decision lines instead of $\overline{\overline{X}}$ and s.

2.3.1 ANOM under Measurement Error

In ANOM, a target line and two action (decision) lines have been drawn one on each side of the target line at a distance of say $h_\alpha \frac{\sigma}{\sqrt{n}}$ (σ is assumed to be known). The observations are plotted on the graph corresponding to sample numbers and an out-of-control alarm is triggered whenever an observation occurs outside the region enclosed by the action lines. A process is in-control when no systematic bias exists, where the bias is defined as the difference between the process mean and the target value.

Suppose now that the observations are subject to some measurement error, i.e., what are observed is $x_i = p_i + e_i$, where p_i is the measured variable and e_i is the measurement error. Let us assume that p_i's are normally distributed with mean μ and variance σ_p^2 and that the measurement errors are distributed independently as $N(\mu_e, \sigma_e^2)$. Furthermore, we assume that the p_i's and e_i's are uncorrelated. Then the decision lines are given by,

$$\begin{aligned} CL &= \mu_1 \\ UDL &= \mu_1 + h_\alpha \frac{\sigma_1}{\sqrt{n}} \\ LDL &= \mu_1 - h_\alpha \frac{\sigma_1}{\sqrt{n}}, \end{aligned} \tag{29}$$

where $\mu = \mu + \mu_e$ and $\sigma_1 = \sqrt{\sigma_p^2 + \sigma_e^2}$.

In the concept of control charts, average run length (ARL) is defined as the expected number of samples to be taken before a false alarm is signaled. In the case of the ANOM with action lines of $h_\alpha \frac{\sigma}{\sqrt{n}}$ units from the target line, the ARL may be given as a fraction of the bias θ (in units of σ_1), where the bias is defined as the difference between the process mean and the target value and is given as $ARL = \frac{1}{1-\beta}$ where $\beta =$ Pr(not detecting this shift on the subsequent sample).

Consider the OC curve for an $\overline{X}$ chart with the standard deviation σ is known. If the mean shifts from the in-control value, say μ_0, to another value $\mu_1 = \mu_0 + c\sigma$, where c is a constant, the probability of not detecting this shift on the first subsequent sample or the β risk is

$$\beta = Pr\{LDL \le \overline{\overline{X}} \le UDL | \mu_i = \mu_1 = \mu_0 + c\sigma\}. \tag{30}$$

Since it is assumed that the i^{th} group mean $\overline{X}_i \sim N(\mu, \frac{\sigma^2}{n})$ then by analogy it can be proved as $\overline{\overline{X}} \sim N(\mu, \frac{\sigma^2}{kn})$. That is, if $\overline{X}_i \sim N(\mu, \frac{\sigma^2}{n})$, then $E(\overline{\overline{X}}) = E(\frac{1}{k}\sum_{i=1}^k \overline{X}_i) = \frac{1}{k}\sum_{i=1}^k E(\overline{X}_i) = \frac{k\mu}{k} = \mu$. Similarly, $Var(\overline{\overline{X}}) = Var(\frac{1}{k}\sum_{i=1}^k \overline{X}_i) = \frac{1}{k^2}\sum_{i=1}^k Var(\overline{X}_i) = \frac{1}{k^2}k(\frac{\sigma^2}{n}) = \frac{\sigma^2}{kn}$.

The upper and lower decision lines for ANOM are given by

$$\begin{aligned} UDL &= \mu_0 + h_\alpha \frac{\sigma}{\sqrt{kn}} \\ LDL &= \mu_0 - h_\alpha \frac{\sigma}{\sqrt{kn}}. \end{aligned}$$

We may write equation (30) as,

$$\begin{aligned}\beta &= \Phi\left\{\frac{[UDL-(\mu_0+c\sigma)]}{\sigma/\sqrt{kn}}\right\} - \Phi\left\{\frac{[LDL-(\mu_0+c\sigma)]}{\sigma/\sqrt{kn}}\right\}\\ &= \Phi\left\{\frac{[\mu_0+h_\alpha(\sigma/\sqrt{kn})-(\mu_0+c\sigma)]}{\sigma/\sqrt{kn}}\right\} - \Phi\left\{\frac{[\mu_0-h_\alpha(\sigma/\sqrt{kn})-(\mu_0+c\sigma)]}{\sigma/\sqrt{kn}}\right\}\\ &= \Phi\left\{h_\alpha - c\sqrt{kn}\right\} - \Phi\left\{-h_\alpha - c\sqrt{kn}\right\} \qquad (31)\end{aligned}$$

where $\Phi(.)$ denotes the standard normal cumulative distribution function. Then the ARL for ANOM is given by

$$ARL = \frac{1}{1-\beta} \qquad (32)$$

where $\beta = \frac{1}{\sqrt{2\pi}}\int\limits_{z_1}^{z_2} e^{\frac{-z^2}{2}}dz,\ z_1 = -h_\alpha - c\sqrt{kn}$ and $z_2 = h_\alpha - c\sqrt{kn}$.

Table 4: ARL When There is No Measurement Error

Actual Bias θ (in units of σ_1)	ARL
0.0	142.9
0.4	98.04
0.8	38.8
1.0	25.0
1.5	9.5

For example, suppose that $n = 3$ and we wish to determine the probability of detecting a shift to $\mu_1 = \mu_0 + 2\sigma$ on the first sample following the shift. Since $c = 2, n = 3$ and $k = 3$ (say), the degrees of freedom is $k(n-1) = 6$ so that we have $h_\alpha = 3.07$ and therefore

$$\begin{aligned}\beta &= \Phi(3.07 - 2*3) - \Phi(-3.07 - 2*3)\\ &= \Phi(-2.93) - \Phi(-9.07) \approx 0.0017.\end{aligned}$$

This is the β-risk or the probability of not detecting such a shift. The probability that such a shift will be detected on the first subsequent sample is $1-\beta = 1 - 0.0017 = 0.9983$ and then $ARL = \frac{1}{1-\beta} = 1.0017$.

Table 5: ARL When There is Measurement Error

Bias θ	$\sigma_e = 0.5$				$\sigma_e = 1.0$			
	Measurement Bias μ_e				Measurement Bias μ_e			
	-0.5	0.0	0.5	1.0	-0.5	0.0	0.5	1.0
0.0	27.1	40.7	27.1	13.0	15.3	19.0	15.0	9.0
0.4	39.5	30.8	15.0	7.4	19.08	16.3	10.0	6.0
0.8	34.2	17.3	8.4	4.5	17.50	11.14	7.0	4.0
1.0	59.2	60.6	6.5	3.7	15.0	9.0	5.3	3.4
1.2	20.04	9.75	5.13	3.0	12.4	7.23	4.34	3.0
1.5	13.0	6.50	4.0	2.4	9.0	5.24	3.4	2.3

To construct the OC curve for the ANOM, plot the β-risk against the magnitude of the shift we wish to detect expressed in standard deviation units for various sample sizes n. These probabilities may be evaluated directly from equation (30). Under measurement error the limits can be written as (see Abraham (1977)),

$$z_1 = \frac{-h_\alpha - c\sqrt{kn}}{\sqrt{(1+\sigma_e^2)}};\ z_2 = \frac{h_\alpha - c\sqrt{kn}}{\sqrt{(1+\sigma_e^2)}}$$

Table 4 gives the ARL values of the ANOM for some specified bias θ when there is no measurement error in the process. Table 5 yields the ARL of ANOM for specified θ, μ_e and σ_e, when measurement error occurs. The effect of the measurement error is quite clear from these computations. When there is no measurement error there is one false alarm in every 143 observations on the average. However, if $\mu_e = 0.5$ and $\sigma_e = 0.5$ then there would be one false alarm in every 27 observations and if $\mu_e = 0.5$ and $\sigma_e = 1.0$ then there would be one in every 15 on the average. Thus the effect of measurement error has been clearly shown. If one knows μ_e and σ_e in advance then the ANOM chart could be modified by adjusting the target value and by recalling the ordinates.

3. Conclusions

In this article we have seen that both types of errors that are not unrealistic in industry seriously affect the OC curve of ANOM under both proportions and count data. In particular, when the center line and the decision limits are based on a target value, the process can easily be judged in-control when, if in fact, it is not. When ANOM chart is based upon data under inspection error, the OC curve of ANOM again distorted but not nearly so seriously. Adjusted decision limits have been presented which do alleviate, to a considerable extent, this problem. We have assumed throughout a constant error. In reality, this error could vary with inspection, shifts, types of nonconforming (nonconformities) etc. Therefore one can try for a separate study for these cases. We have also shown the impact of measurement error on the ANOM chart. When there is no measurement error, the rate of false alarm is sufficiently reduced.

Acknowledgements
The authors are thankful to the editor and the reviewers for their constructive comments on this paper.

References

[1] B. Abraham, *Control Chart and Measurement Error*, Proceedings of 31st Annual Technical Conference, American Society for Quality Control (ASQC), (1977), Philadelphia.

[2] S. Balamurali, Simplified Factors of Analysis of Means for Treatment Effects, *Int. J. Stat. Sys.*, **5**(2010), 135-145.

[3] L. S. Nelson, Exact Critical Points for the Analysis of Means, *J. Qual.Techn.*, **15**(1983a), 40-44.

[4] L. S. Nelson, Transformations for Attribute Data, *J. Qual. Techn.*, **15**(1983b), 45-54.

[5] E. R. Ott, E. R, Analysis of Means - A Graphical Procedure, *Ind. Qual. Control*, **24**(1967), 101-109.

[6] E. R. Ott, *Process Quality Control*, New York, McGraw-Hill, 1975.

[7] P. F. Ramig, Applications of the Analysis of Means, *J. Qual. Techn.* **15**(1883), 19-25.

[8] Ron Suich, The c Chart under Inspection Error, *J. Qual. Techn.,* **20**(1988), 263-266.

[9] E. G. Schilling, A Systematic Approach to the Analysis of Means - Part I. Analysis of Treatment Effects, *J. Qual. Techn.,* **5**(1973), 93-108.

[10] E. G. Schilling, A Systematic Approach to the Analysis of Means - Part II. Analysis of Contrasts. Part III: Analysis of Non-normal Data, *J. Qual. Techn*, **5**(1973), 147-159.

[11] H. M. Wadsworth, K. S. Stephens, and A. B. Godfrey, A.B, *Modern Methods for Quality Control and Improvement*, John Wiley and sons, Yew York (1986).