IEOR 165 – Lecture 22
Weighted Control Charts

1 Other Control Charts

1.1 Fraction Defective

Suppose that when the process is in control, the probability that a single item will be defective is p. Now suppose $X_1, X_2, \dots$ is zero (with probability $1-p$) if the corresponding item is not defective, and is one (with probability p) if the corresponding item is defective. The main idea is that if we look at the sum of a subgroup

$$V^{j+1} = \sum_{i=1}^{n} X_{jn+i},$$

then this quantity has a binomial distribution with n trials and success probability p. Because the expected number of defects in one subgroup is np, we need n to be fairly large in order to observe some defects (since p will typically be relatively small). However, when n is large, we can approximate a binomial distribution by $\mathcal{N}(np, np(1-p))$.
This approximation can be used to construct a control chart. In particular, if we define

$$F^{j+1} = \frac{1}{n}V^{j+1},$$

then this is approximately distributed as $\mathcal{N}(p, p(1-p)/n)$. And so we can construct the control chart for F^{j+1} using the confidence intervals for a Gaussian distribution. In particular, we would use

$$\begin{aligned} LCL &= p - z(1-\alpha/2) \cdot \sqrt{\frac{p(1-p)}{n}} \\ UCL &= p + z(1-\alpha/2) \cdot \sqrt{\frac{p(1-p)}{n}}. \end{aligned}$$

And we say the process is out of control if $F^{j+1} \notin [LCL, UCL]$.

1.2 Number of Defects

Suppose the number of defects in the i-th item X_i is described by a Poisson random variable with mean λ. One example where this occurs is when there are K different possible faults, and each fault can occur with probability p. If K is large relative to p, then the binomial distribution with K trials and success probability p can be approximated by a Poission distribution with mean Kp. As a result, when we talk about the number of defects we will exam one item at a time.

Let L be a Poission random variable with mean 1, and define $\ell(1-\alpha)$ to be the value θ such that $\mathbb{P}(L \leq \theta) = 1-\alpha$. Then the confidence interval is

$$\mathbb{P}\left(\ell(\alpha/2) \leq \frac{X_i}{\lambda} \leq \ell(1-\alpha/2)\right) = 1-\alpha$$
$$\Rightarrow \underline{\mu}^i = \frac{X_i}{\ell(1-\alpha/2)}, \qquad \overline{\mu}^i = \frac{X_i}{\ell(\alpha/2)}.$$

Consequently, we have that

$$LCL = \ell(\alpha/2) \cdot \lambda$$
$$UCL = \ell(1-\alpha/2) \cdot \lambda,$$

and we say the process is out of control if $X_i \notin [LCL, UCL]$.

2 Changes in Population Mean

Another situation in which control charts are used is when we are interested in detecting changes in the mean of a process. There are multiple possible approaches.

2.1 Moving-Average Control Charts

Suppose $X_1, X_2, \ldots \sim \mathcal{N}(\mu, \sigma^2)$ are measurements, and let $\overline{X}^k$ be the sample average of the k-th subgroup of data (of size n points). Define the moving average as

$$M^k = \begin{cases} \frac{1}{m} \cdot (\overline{X}^{k-m+1} + \ldots + \overline{X}^{k-1} + \overline{X}^k), & \text{if } k \geq m \\ \frac{1}{k} \cdot (\overline{X}^1 + \ldots + \overline{X}^k), & \text{otherwise} \end{cases}$$

This is a moving average over m steps. Because these quantities are simply linear combinations of Gaussians, we can simply derive the upper control limits:

$$UCL^k = \begin{cases} \mu + \sigma \cdot z(1-\alpha/2)/\sqrt{nk}, & \text{if } k < m \\ \mu + \sigma \cdot z(1-\alpha/2)/\sqrt{nm}, & \text{otherwise} \end{cases}$$

and the lower control limits:

$$LCL^k = \begin{cases} \mu - \sigma \cdot z(1-\alpha/2)/\sqrt{nk}, & \text{if } k < m \\ \mu - \sigma \cdot z(1-\alpha/2)/\sqrt{nm}, & \text{otherwise} \end{cases}$$

We say that the process is out of control if M^k exceeds the limits defined by LCL^k/MCL^k.

2.2 Exponentially Weighted Moving-Average Control Charts

Another weighting scheme that is possible is a recursive weighting

$$M^k = \gamma \cdot \overline{X}^k + (1-\gamma) \cdot M^{k-1},$$

where $\gamma \in (0, 1]$ is a constant. The M^k is again a linear combination of Gaussians, and for k large the M^k approximately have a variance of $\sigma^2\gamma/(n(2-\gamma))$. Thus, the control limits are given by

$$\begin{aligned} LCL^k &= \mu - \sigma \cdot z(1-\alpha/2)/\sqrt{n(2-\gamma)/\gamma} \\ UCL^k &= \mu + \sigma \cdot z(1-\alpha/2)/\sqrt{n(2-\gamma)/\gamma}. \end{aligned}$$

2.3 Choosing the Significance Level

The equation we derived in the previous lecture relating the significance level to the expected number of subgroups that would be analyzed before the process is declared to be out of control under the situation where the process was never out of control does not directly apply to this case because it assumes independence between the different hypothesis tests. This independence assumption is clearly not true for control charts using moving-averages, and so the question is how should we instead choose the significance level for control charts based on moving-averages? Instead of using an equation, an alternative approach is to use a Monte Carlo algorithm. In particular, suppose we have a set $A = \{\alpha_1, \ldots, \alpha_m\}$ of possible values of the significance level, and let M be a given large value (say $M = 1000$). Then we can do the following

1. For each value $\alpha_k \in A$
   (a) For $ind = 1, \ldots, M$
      i. Set $j = 0$
      ii. Do
         A. Randomly choose $X_{jn+i} \sim \mathcal{N}(\mu, \sigma^2)$ for $i = 1, \ldots, n$.
         B. Compute M^{j+1} using one of the moving-average formulas.
         C. Set $j = j + 1$.
      iii. While $M^j \in [LCL^j, UCL^j]$, where the equations for LCL^j/UCL^j correspond to the particular moving-average formula used.
      iv. Set $E_{ind} = j$.
   (b) Set $L_k = \frac{1}{M}\sum_{ind=1}^{M} E_{ind}$

The result of this algorithm will be a a set of pairs of values (α_k, L_k), which relates a given significance level α_k to the expected number of subgroups analyzed L_k. This is a Monte Carlo algorithm because we approximate the expectations by sampling many random values and using sample averages as the approximation of the expectations.