Ecological Archives A014-036-A1

Marti J. Anderson and Angus A. Thompson. 2004. Multivariate control charts for ecological and environmental monitoring. *Ecological Applications* 14:1921-1935.

Appendix A. Derivation of a multivariate control chart criterion using interpoint distances.

Let x_{ik} be the ith observation in a set of $i = 1, \ldots, t$ points on variable k for $k = 1, \ldots, p$ variables and let $\bar{x}_{tk} = \frac{1}{t}\sum_{i=1}^{t} x_{ik}$ and $\bar{x}_{(t-1)k} = \frac{1}{t-1}\sum_{i=1}^{t-1} x_{ik}$. We wish to obtain the following criterion: $d_t = \sqrt{\sum_{k=1}^{p}(x_{tk} - \bar{x}_{(t-1)k})^2}$ using only inter-point distances.

First, define two quantities which can be obtained using only inter-point distances: the sum of squared distances from the points to their centroid for the first (t - 1) points:

$$SS_{t-1} = \sum_{k=1}^{p}\sum_{i=1}^{t-1}(x_{ik} - \bar{x}_{(t-1)k})^2 = \frac{1}{(t-1)}\sum_{p=1}^{k}\sum_{i=1}^{t-2}\sum_{j=i+1}^{t-1}(x_{ik} - x_{jk})^2 \tag{A.1}$$

and for all t points:

$$SS_t = \sum_{k=1}^{p}\sum_{i=1}^{t}(x_{ik} - \bar{x}_{tk})^2 = \frac{1}{t}\sum_{p=1}^{k}\sum_{i=1}^{t-1}\sum_{j=i+1}^{t}(x_{ik} - x_{jk})^2 \; . \tag{A.2}$$

Now, we can write:

$$SS_t = \sum_{k=1}^{p}\sum_{i=1}^{t}[(x_{ik} - \bar{x}_{(t-1)k}) + (\bar{x}_{(t-1)k} - \bar{x}_{tk})]^2 \tag{A.3}$$

and expanding the square this gives:

$$SS_t = \sum_{k=1}^{p}\left\{\sum_{i=1}^{t}(x_{ik} - \bar{x}_{(t-1)k})^2 + t(\bar{x}_{(t-1)k} - \bar{x}_{tk})^2 + 2(\bar{x}_{(t-1)k} - \bar{x}_{tk})\sum_{i=1}^{t}(x_{ik} - \bar{x}_{(t-1)k})\right\}. \tag{A.4}$$

Recognizing the first term in the *rhs* of the expression, we have:

$$SS_t = SS_{t-1} + \sum_{k=1}^{p}\Big\{(x_{tk} - \bar{x}_{(t-1)k})^2 + t(\bar{x}_{(t-1)k} - \bar{x}_{tk})^2 + 2(\bar{x}_{(t-1)k} - \bar{x}_{tk})\sum_{i=1}^{t}(x_{ik} - \bar{x}_{(t-1)k})\Big\}. \tag{A.5}$$

Next, we can simplify the following expression $(\bar{x}_{(t-1)k} - \bar{x}_{tk})$ to

$$\frac{\sum_{i=1}^{t-1} x_{ik}}{(t-1)} - \frac{\sum_{i=1}^{t} x_{ik}}{t} = \frac{\sum_{i=1}^{t-1} x_{ik}}{t(t-1)} - \frac{x_{tk}}{t} = \frac{(\bar{x}_{(t-1)k} - x_{tk})}{t}, \tag{A.6}$$

and plugging this quantity where appropriate into Eq. A.5, we have

$$SS_t = SS_{t-1} + \sum_{k=1}^{p}(x_{tk} - \bar{x}_{(t-1)k})^2\left[1 + \tfrac{1}{t} - \tfrac{2}{t}\right] \tag{A.7}$$

which, rearranging, gives

$$d_t^2 = \tfrac{t}{t-1}(SS_t - SS_{t-1}) \tag{A.8}$$

Q.E.D.