Supplemental Material for "Online and Non-parametric Drift Detection Methods Based on Hoeffding's Bounds"

Isvani Frías-Blanco, José del Campo-Ávila, Gonzalo Ramos-Jiménez, Rafael Morales-Bueno, Agustín Ortiz-Díaz and Yailé Caballero-Mota

Appendix A Proofs

A.1 Proof of Corollary 3

Let us first consider the probabilistic bound for the type I error (probability of false detection). If $E\left[\overline{X}\right] \leq E\left[\overline{Y}\right]$ (or equivalently $E\left[\overline{X}\right] - E\left[\overline{Y}\right] \leq 0$), by inclusion (i.e. if $A \Rightarrow B$ then $\Pr(A) \leq \Pr(B)$) we have ($D = \overline{X} - \overline{Y}$ and $E_D = E\left[\overline{X}\right] - E\left[\overline{Y}\right]$)

$$\Pr\{D \geq \varepsilon_\alpha\} \leq \Pr\{D - E_D \geq \varepsilon_\alpha\} \tag{11}$$

Applying inversion on equation (1):

$$\Pr\{D - E_D \geq \varepsilon_\alpha\} \leq \alpha \tag{12}$$

Applying transitivity on equations (**??**) and (**??**), we obtain $\Pr\left\{\overline{X} - \overline{Y} \geq \varepsilon_\alpha\right\} \leq \alpha$ as desired.

Consider now the bound for the type II error (probability of non-detection). If $E\left[\overline{X}\right] \geq E\left[\overline{Y}\right] + \varsigma$ then for $\varsigma > \varepsilon_\alpha$

$$\begin{aligned}\Pr\{D < \varepsilon_\alpha\} &= \Pr\{D - \varsigma < \varepsilon_\alpha - \varsigma\} \\ &\leq \Pr\{D - E_D < \varepsilon_\alpha - \varsigma\}\end{aligned} \tag{13}$$

Since for $\varepsilon > 0$, $\Pr\{D - E_D > \varepsilon\} = \Pr\{D - E_D < -\varepsilon\}$; as $\varepsilon_\alpha - \varsigma < 0$, from equation (1):

$$\Pr\{D - E_D < \varepsilon_\alpha - \varsigma\} \leq e^{\frac{-2(\varsigma - \varepsilon_\alpha)^2}{(n^{-1} + m^{-1})(b-a)^2}} \tag{14}$$

Finally, again for transitivity on equations (**??**) and (**??**), we have

$$\Pr\left\{\overline{X} - \overline{Y} < \varepsilon_\alpha\right\} \leq e^{\frac{-2(\varsigma - \varepsilon_\alpha)^2}{(n^{-1} + m^{-1})(b-a)^2}}$$

A.2 Proof of Corollary 5

Let $\overrightarrow{X} = (X_1, X_2, \ldots, X_n, Y_1, Y_2, \ldots, Y_m)$ and let $f(\overrightarrow{X}) = \hat{X}_n - \hat{Y}_m$. Then, if $\hat{X}_n$ and $\hat{Y}_n$ satisfy the independent bounded differences condition on Theorem 4, it is easy to prove that for $f(\overrightarrow{X})$ the independent bounded differences condition is satisfied taking $d_i = (b-a)v_i$ for all $i \in \{1, 2, \ldots, n\}$, and $d_j = (b-a)v_j^{'}$ for all $j \in \{1, 2, \ldots, m\}$. So

$$\sum_{i=1}^{n+m} d_i^2 = (b-a)^2 \left[\sum_{i=1}^{n} (v_i)^2 + \sum_{i=1}^{m} \left(v_i^{'}\right)^2\right] = \mathcal{D}_{n,m}.$$

Finally, taking all together we obtain the desired bound from equation (3).

A.3 Proof of Example 7

At n random variables, the weights of the EWMA statistic take the form $v_1 = (1-\lambda)^{n-1}$, and $v_i = \lambda(1-\lambda)^{n-i}$ for $1 < i \leq n$ (note that $\sum_{i=1}^{n} v_i = 1$). Then for $n > 1$ ($\mathcal{D}_1 = 1$) we have:

$$\begin{aligned}\mathcal{D}_n &= \sum_{i=1}^{n} d_i^2 = (b-a)^2 \sum_{i=1}^{n} v_i^2 \\ &= (b-a)^2 \left[\sum_{i=2}^{n} \lambda^2 (1-\lambda)^{2(n-i)} + (1-\lambda)^{2(n-1)}\right] \\ &= (b-a)^2 \left[\lambda^2 \frac{1 - (1-\lambda)^{2(n-1)}}{1 - (1-\lambda)^2} + (1-\lambda)^{2(n-1)}\right]\end{aligned}$$

When $n \to \infty$, $\mathcal{D}_n$ converge to a constant value as $\lim_{n\to\infty} \mathcal{D}_n = (b-a)^2 \lambda/(2-\lambda)$. Since $\sum_{i=1}^{n} d_i^2 = \mathcal{D}_n$, $f(\overrightarrow{X}) = \hat{X}_n$ and $E[f(\overrightarrow{X})] = E\left[\hat{X}_n\right]$; from equation (3) we obtain the desired bound.

A.4 Proof of Example 8

This time, the independent bounded differences condition is

$$\begin{aligned}\mathcal{D}_n &= \sum_{i=1}^{n} d_i^2 = \left(\frac{b-a}{\mathcal{W}_n}\right)^2 \sum_{i=1}^{n} \beta^{2(n-i)} \\ &= (b-a)^2 \left(\frac{1-\beta}{1-\beta^n}\right)^2 \left(\frac{1-\beta^{2n}}{1-\beta^2}\right)\end{aligned}$$

As $0 < \beta \leq 1$, $\mathcal{D}_n$ converge to $\lim_{n\to\infty} \mathcal{D}_n = (b-a)^2 (1-\beta)/(1+\beta)$. As above, from equation (3) we can obtain the desired bound.