Errata "Algorithmic Luckiness"

15th April 2004

1 Proof of Lemma 20

The article *R. Herbrich and R. Williamson. Algorithmic Luckiness. Journal of Machine Learning Research 3. pp. 175-212. 2002* contains a mistake on page 195. In the proof of Lemma 20 it is argued that the probability that a binomially distributed random variable with an expectation of more than ε is greater than or equal to $\frac{\varepsilon(n-m)}{2}$ is *at least* $1-(1-\varepsilon)^{n-m}$ provided $\varepsilon(n-m) \geq 2$.

This is wrong; in order to see this let A and B be defined as follows

$$A := \left\{ i \in \mathbb{N} \,\middle|\, i \geq \frac{\varepsilon(n-m)}{2} \right\}, \quad B := \{ i \in \mathbb{N} \,|\, i \geq 1 \} .$$

Since $\varepsilon(n-m) \geq 2$ we know that $A \subseteq B$ and thus $\mathsf{P}(A) \leq \mathsf{P}(B)$. By the binomial tail bound we know that $\mathsf{P}(\overline{B}) \leq (1-\varepsilon)^{n-m}$ and thus $\mathsf{P}(B) \geq 1-(1-\varepsilon)^{n-m}$. Now we can see that the paper incorrectly tied a lower bound on $\mathsf{P}(B)$ with an upper bound on $\mathsf{P}(B)$.

Nevertheless, the lemma remains true if we use the following theorem due to Mingrui Wu. In the current application we replace n in the theorem with $n-m$ from Lemma 20 and μ in the theorem with ε from Lemma 20.

Theorem (Binomial mean deviation bound). *Let $\mathsf{X}_1, \ldots, \mathsf{X}_n$ be independent random variables such that, for all $i \in \{1, \ldots, n\}$, $\mathsf{P}_{\mathsf{X}_i}(\mathsf{X}_i = 1) = 1 - \mathsf{P}_{\mathsf{X}_i}(\mathsf{X}_i = 0) = \mathsf{E}_{\mathsf{X}_i}[\mathsf{X}_i] = \mu$. Then, for all $\varepsilon \in \left(\frac{2}{n}, \mu\right)$ we have*

$$\mathsf{P}_{\mathsf{X}^n}\left(\frac{1}{n} \sum_{i=1}^{n} \mathsf{X}_i \geq \frac{\varepsilon}{2} \right) > \frac{1}{2} .$$

Proof. Since $\mu > \varepsilon$ it suffices to show

$$\mathsf{P}_{\mathsf{X}^n}\left(\frac{1}{n} \sum_{i=1}^{n} \mathsf{X}_i \geq \frac{\varepsilon}{2} \right) \geq \mathsf{P}_{\mathsf{X}^n}\left(\frac{1}{n} \sum_{i=1}^{n} \mathsf{X}_i \geq \frac{\mu}{2} \right) > \frac{1}{2} ,$$

assuming that $n\mu > 2$. This statement is equivalent to

$$\mathsf{P}_{\mathsf{X}^n}\left(\sum_{i=1}^{n} \mathsf{X}_i < \frac{n\mu}{2} \right) \leq \frac{1}{2} . \tag{1.1}$$

Let l be the largest integer such that $l < \frac{n\mu}{2}$. Since $\mu \in [0,1]$ and n is an integer we know that $2l+1 \leq n$. Note that $\mathsf{S} := \sum_{i=1}^{n} \mathsf{X}_i$ is binomially distributed with parameters n and μ. Thus, (1.1) is equivalent to

$$\sum_{j=0}^{l} \binom{n}{j} \mu^j (1-\mu)^{n-j} \leq \frac{1}{2} . \tag{1.2}$$

Case 1: $\mu > \frac{1}{2}$ In this case $\mu > 1-\mu$ and for $j \in \{0, \ldots, l\}$ we have $j < n-j$ so it follows that

$$\binom{n}{j} \mu^j (1-\mu)^{n-j} < \binom{n}{j} \mu^{n-j} (1-\mu)^j = \binom{n}{n-j} \mu^{n-j} (1-\mu)^j .$$

Hence, double summation of (1.2) gives

$$\begin{aligned} 2\sum_{j=0}^{l}\binom{n}{j}\mu^j(1-\mu)^{n-j} &< \sum_{j=0}^{l}\binom{n}{j}\mu^j(1-\mu)^{n-j}+\sum_{j=n-l}^{n}\binom{n}{j}\mu^j(1-\mu)^{n-j} \\ &\le \sum_{j=0}^{n}\binom{n}{j}\mu^j(1-\mu)^{n-j}=1\,,. \end{aligned}$$

Case 2: $\mu \le \frac{1}{2}$ By assumption $n\mu > 2$ and thus $l \le \frac{n}{4}$ and $n > 4$. In the rest of the proof we will show that

$$\forall j \in \{1,\ldots,l\}: \quad \binom{n}{j}\mu^j(1-\mu)^{n-j} < \binom{n}{j+l}\mu^{j+l}(1-\mu)^{n-j-l}\,, \tag{1.3}$$

$$(1-\mu)^n < \binom{n}{2l+1}\mu^{2l+1}(1-\mu)^{n-2l-1}\,. \tag{1.4}$$

Using these two results, (1.2) can be seen to hold by noticing that (1.3) and (1.4) imply

$$\begin{aligned} \sum_{j=0}^{l}\binom{n}{j}\mu^j(1-\mu)^{n-j} &= \sum_{j=1}^{l}\binom{n}{j}\mu^j(1-\mu)^{n-j}+(1-\mu)^n \\ &< \sum_{j=l+1}^{2l+1}\binom{n}{j}\mu^j(1-\mu)^{n-j}\,. \end{aligned}$$

Hence, double summation of (1.2) again gives

$$\begin{aligned} 2\sum_{j=0}^{l}\binom{n}{j}\mu^j(1-\mu)^{n-j} &< \sum_{j=0}^{2l+1}\binom{n}{j}\mu^j(1-\mu)^{n-j} \\ &\le \sum_{j=0}^{n}\binom{n}{j}\mu^j(1-\mu)^{n-j}=1\,, \end{aligned}$$

where we used the fact that $2l+1 \le n$. It remains to show (1.3) and (1.4). In order to prove (1.3) we divide the right hand side by the left hand side. For the jth term this results in

$$\begin{aligned} \frac{\binom{n}{j+l}\mu^{j+l}(1-\mu)^{n-j-l}}{\binom{n}{j}\mu^j(1-\mu)^{n-j}} &= \prod_{t=1}^{l}\frac{\mu}{1-\mu}\cdot\frac{n-j-l+t}{j+t} \\ &\ge \prod_{t=1}^{l}\frac{\mu}{1-\mu}\cdot\frac{n-2l+t}{l+t} \\ &= \prod_{t=1}^{l}\frac{\mu}{1-\mu}\left(1+\frac{n-3l}{l+t}\right) \\ &\ge \prod_{t=1}^{l}\frac{\mu}{1-\mu}\cdot\frac{n-l}{2l} \\ &= \left(\frac{\mu}{1-\mu}\cdot\frac{n-l}{2l}\right)^l \\ &> \left(\frac{\mu}{1-\mu}.\frac{n-\frac{n\mu}{2}}{n\mu}\right)^l \\ &= \left(\frac{1-\frac{\mu}{2}}{1-\mu}\right)^l > 1\,, \end{aligned}$$

where we used $j \leq l$ in the second line, $t \leq l$ and $n - 3l \geq 0$ in the third line and $l < \frac{n\mu}{2}$ in the penultimate line. In order to show (1.4) we assume $l \geq 1$; otherwise the statement follows easily. Again, dividing the right hand side of (1.4) by the left hand side of (1.4) we obtain

$$\begin{aligned}
&\frac{\binom{n}{2l+1}\mu^{2l+1}(1-\mu)^{n-2l-1}}{(1-\mu)^n} \\
&\quad = \prod_{t=1}^{2l+1} \frac{\mu}{1-\mu} \cdot \frac{n-2l-1+t}{t} \\
&\quad = \left(\prod_{t=2}^{2l} \frac{\mu}{1-\mu} \cdot \frac{n-2l-1+t}{t}\right)\left(\frac{n(n-2l)}{2l+1}\left(\frac{\mu}{1-\mu}\right)^2\right) \\
&\quad > \left(\prod_{t=2}^{2l} \frac{\mu}{1-\mu} \cdot \frac{n-1}{2l}\right)\left(\frac{n(n-n\mu)}{2l+1}\left(\frac{\mu}{1-\mu}\right)^2\right) \\
&\quad = \left(\frac{n\mu-\mu}{2l-2l\mu}\right)^{2l-1}\left(\frac{n^2\mu^2}{(2l+1)(1-\mu)}\right) \\
&\quad > \left(\frac{2l-\mu}{2l-2l\mu}\right)^{2l-1}\left(\frac{n^2\mu^2}{2l+1}\right) \\
&\quad > \left(\frac{2l-\mu}{2l-2l\mu}\right)^{2l-1}\left(\frac{n^2\mu^2}{n\mu+1}\right) > 1\,,
\end{aligned}$$

where the third and fifth line uses $t \leq 2l < n\mu$ and the last line uses $n\mu > 2$.

The theorem is proven. □