Towards Practical Mean Bounds for Small Samples
Erratum

My Phan [*][†]

November 2, 2021

Lemma 2.2 in our paper [Phan et al., 2021] has the following proof:

Lemma 2.2. *Let X be a random variable with CDF F and $Y \stackrel{\text{def}}{=} F(X)$, known as the probability integral transform of X. Let U be a uniform random variable on $[0,1]$. Then for any $0 \leq y \leq 1$,*

$$\mathbb{P}(Y \leq y) \leq \mathbb{P}(U \leq y). \tag{1}$$

If F is continuous, then Y is uniformly distributed on $(0,1)$.

Proof. Since $\mathbb{P}(U \leq y) = y$, we will show that $\mathbb{P}(Y \leq y) \leq y$.

We will first show that if $F(x) \leq y$, then $x \leq \sup\{z : F(z) \leq y\}$. Suppose that $x > \sup\{z : F(z) \leq y\}$. Then, $F(x) > y$. Therefore,

$$F(x) \leq y \text{ implies } x \leq \sup\{z : F(z) \leq y\}. \tag{2}$$

Now we have

$$\begin{aligned}
\mathbb{P}(Y \leq y) &= \mathbb{P}(F(X) \leq y) &(3)\\
&\leq \mathbb{P}(X \leq \sup\{z : F(z) \leq y\}) &(4)\\
&= F(z^*) \text{ where } z^* = \sup\{z : F(z) \leq y\} &(5)\\
&\leq y. &(6)
\end{aligned}$$

If F is continuous, Angus [1994] shows that Y is uniformly distributed on $(0,1)$. □

The proof is incorrect because the step from Eq. 5 to Eq. 6 $F(z^*) \leq y$ is not true if F is not left-continuous. The corrected proof should be:

Proof. Let $F^{-1}(y) = \inf\{x : F(x) \geq y\}$ for $0 < y < 1$ and U be an uniform random variable on $(0,1)$. Since F is non-decreasing and right-continuous, $F(F^{-1}(y)) \geq y$. By Angus [1994], $F^{-1}(U)$ has CDF F. For $0 < y < 1$, then:

$$\begin{aligned}
\mathbb{P}(Y \leq y) &= \mathbb{P}(F(X) \leq y) &(7)\\
&= \mathbb{P}(F(F^{-1}(U)) \leq y) &(8)\\
&\leq \mathbb{P}(U \leq y) &(9)\\
&= y. &(10)
\end{aligned}$$

If F is continuous, Angus [1994] shows that Y is uniformly distributed on $(0,1)$. □

References

John E. Angus. The probability integral transform and related results. *SIAM Rev.*, 36(4):652–654, December 1994. ISSN 0036-1445. doi: 10.1137/1036146. URL https://doi.org/10.1137/1036146.

My Phan, Philip S. Thomas, and Erik Learned-Miller. Towards practical mean bounds for small samples. In *Proceedings of the 38th International Conference on Machine Learning (ICML-21)*, 2021.

[*]College of Information and Computer Sciences, University of Massachusetts Amherst.
[†]myphan@cs.umass.edu.