Biometrika (2022), **109**, 4, *pp.* 1181–1182 https://doi.org/10.1093/biomet/asac043
Advance Access publication 13 September 2022

Correction to: 'Valid sequential inference on probability forecast performance'

ALEXANDER HENZI and JOHANNA F. ZIEGEL

Institute of Mathematical Statistics and Actuarial Science, University of Bern, Alpeneggstrasse 22, 3012 Bern, Switzerland
alexander.henzi@stat.unibe.ch johanna.ziegel@stat.unibe.ch

Biometrika (2021), **109**, pp. 647–63.

In the paragraph before Proposition 2 of Henzi & Ziegel (2021) it is claimed that

$$e_T = \frac{1}{h}\sum_{k=1}^{h}\prod_{l\in I_k} E_{p_l,\, q_l;\lambda_l}(Y_{l+h}), \quad I_k = \{k + hs\colon s = 0, \dots \lfloor (T-k)/h \rfloor - 1\},$$

is a nonnegative supermartingale. This is generally not true for $h > 1$. As a consequence, Proposition 2 is not correct for $h > 1$, and should be adapted as follows.

PROPOSITION 2. *Let $\tau \in \mathbb{N}$ be a stopping time. Then under the assumptions of Proposition 1,*

$$\mathbb{E}_{\mathbb{Q}}(e_{\tau+h-1}) \leqslant 1, \quad \mathbb{Q} \in \mathcal{H}_{\mathrm{S}}.$$

The quantity $p_{t_0} = \min\{1, \inf_{s=1,\dots,t_0} 1/e_s\}$ defined in the last paragraph of § 3 is an anytime-valid p-value only for $h = 1$, but the stopping time $\tau_{\alpha,h}$ guarantees that $\mathbb{Q}(\tau_{\alpha,h} < \infty) \leqslant \alpha$ for $h > 1$, $\alpha \in (0,1)$, and $\mathbb{Q} \in \mathcal{H}_{\mathrm{S}}$, because $\tau_{\alpha,h} < \infty$ implies $e_{\tau_{\alpha,h}+h-1} \geqslant 1/\alpha$. Hence all empirical results in the article remain valid. An anytime-valid p-value for $h > 1$ is given by

$$p_{t_0} = \min\left(1, \inf_{s=1,\dots,t_0}\left[\max_{j=s-h+1,\dots,s-1} E_{p_j,q_j;\lambda_j}\{\mathbb{1}(p_j > q_j)\}^{-1}/e_s\right]\right),$$

since $p_t \leqslant \alpha$ for some $t \in \mathbb{N}$ if and only if $\tau_{\alpha,h} < \infty$.

Proof of Proposition 2. Recall that the process $(Y_t, p_t, q_t, \lambda_t)_{t\in\mathbb{N}}$ is adapted to $\mathfrak{F} = (\mathcal{F}_t)_{t\in\mathbb{N}}$. Let $h > 1$. For $k = 1, \dots, h$, define $I_k(t) = \{k + hs\colon s = 0, \dots \lfloor (t-k)/h \rfloor - 1\}$,

$$M_t^{[k]} = \prod_{l\in I_k(t)} E_{p_l,q_l;\lambda_l}(Y_{l+h}), \quad \mathfrak{F}^{[k]} = \left(\mathcal{F}_{\lfloor \frac{t-k}{h} \rfloor h+k}\right)_{t\in\mathbb{N}},$$

with $\prod_{\emptyset} := 1$ and $\mathcal{F}_j := \{\Omega, \emptyset\}$ for $j \leqslant 0$. Then $e_t = \sum_{k=1}^{h} M_t^{[k]}/h$. For $k = 1, \dots, h$, the process $(M_t^{[k]})_{t\in\mathbb{N}}$ is a nonnegative supermartingale with respect to $\mathfrak{F}^{[k]}$ for any $\mathbb{Q} \in \mathcal{H}_{\mathrm{S}}$, and therefore satisfies $\mathbb{E}_{\mathbb{Q}}(M_{\tau[k]}^{[k]}) \leqslant 1$ for any $\mathfrak{F}^{[k]}$-stopping time $\tau^{[k]}$. So

$$\mathbb{E}_{\mathbb{Q}}\left(\frac{1}{h}\sum_{\ell=1}^{h} M_{\tau^{[k]}}^{[k]}\right) \leqslant 1.$$

for $\mathfrak{F}^{[k]}$-stopping times $\tau^{[k]}$, $k = 1, \ldots, h$. If τ is an $\mathfrak{F}$-stopping time, then

$$\left(\left\lfloor \frac{\tau - k - 1}{h} \right\rfloor + 1 \right) h + k =: f_k(\tau) \in \{\tau, \ldots, \tau + h - 1\}$$

is an $\mathfrak{F}^{[k]}$-stopping time. To see this, let $t = k + hs + j$ for $s \in \mathbb{N}_0$, $k \in \{1, \ldots, k\}$, $j \in \{0, \ldots, h-1\}$. Then $\lfloor (t-k)/h \rfloor h + k = k + hs$, and $f_k(\tau) \leqslant t$ if and only if $\tau \leqslant k + hs$, so

$$\{f_k(\tau) \leqslant t\} = \{\tau \leqslant k + hs\} \in \mathcal{F}_{k+hs} = \mathcal{F}_{\lfloor \frac{t-k}{h} \rfloor h + k}.$$

This implies that for any $\mathfrak{F}$-stopping time τ, we obtain

$$\mathbb{E}_{\mathbb{Q}}(M_{\tau+h-1}) = \mathbb{E}_{\mathbb{Q}} \left(\frac{1}{h} \sum_{k=1}^{h} M^{[k]}_{f_k(\tau)} \right) \leqslant 1, \quad \mathbb{Q} \in \mathcal{H}_{\mathrm{S}},$$

using the fact that $M_{t+h-1} = \sum_{k=1}^{h} M^{[k]}_{f_k(t)}/h$ for $t \in \mathbb{N}$. □

The following example demonstrates that the statement of Proposition 2 with τ instead of $\tau + h - 1$ is not true. Let $h = 2$, $\varepsilon \in (0, 1)$, and $\delta \in (0, \varepsilon)$. Define $p_1 = \varepsilon - \delta$, $q_1 = \varepsilon + \delta$, and $p_t = q_t = 0.5$ for $t > 1$. Let $S(p, y) = (p - y)^2$. Then the one-period e-value for $t = 1$ equals

$$E^{\pi_{1,1}}_{p_1,q_1}(y) = \frac{\pi_{1,1}^{y}(1 - \pi_{1,1})^{1-y}}{\varepsilon^{y}(1-\varepsilon)^{1-y}},$$

with $\pi_{1,1} \in (\varepsilon, 1]$, and $E^{\pi_{1,t}}_{p_t,q_t}(y) \equiv 1$ for $t > 1$, which gives

$$e_t = \begin{cases} 1, & t = 1, 2, \\ 0.5 E^{\pi_{1,1}}_{p_1,q_1}(Y_3) + 0.5, & t \geqslant 3. \end{cases}$$

The null hypothesis consists of all distributions $\mathbb{Q}$ generating the sequence $(Y_t)_{t \in \mathbb{N}}$ such that $\mathbb{Q}(Y_3 = 1 \mid Y_1) \leqslant \varepsilon$. Define $\mathbb{Q}$ as follows. Let $\mathbb{Q}(Y_1 = 1)$ be arbitrary; Y_2, Y_3 independent of Y_1 with $Y_2 = Y_3$ almost surely and $\mathbb{Q}(Y_2 = Y_3 = 1) = p \in (0, \varepsilon]$; and Y_t for $t > 3$ with arbitrary distribution. Define the stopping time $\tau = 3Y_2 + 2(1 - Y_2)$. Then,

$$e_\tau = \begin{cases} 1, & Y_2 = 0, \\ 0.5\pi_{1,1}/\varepsilon + 0.5, & Y_2 = 1, \end{cases}$$

so that $\mathbb{E}_{\mathbb{Q}}(e_\tau) > 1$ for $\pi_{1,1} > \varepsilon$, since $\mathbb{Q}(Y_2 = 1) = p > 0$.

REFERENCES

HENZI, A. & ZIEGEL, J. F. (2021). Valid sequential inference on probability forecast performance. *Biometrika* **109**, 647–63.

[*Received on* 1 *July* 2022. *Editorial decision on* 1 *July* 2022]