Supplemental Materials

Proof of Proposition 1

Proof. Let us write X as $X = \text{sech}T$ where $T \sim W(\alpha, \beta)$, the cdf of X can be determined as

$$F(x, \alpha, \beta) = P(X \leq x) = P(\text{sech}T \leq x) = P(\text{arcsech}x \leq T \leq \infty) = 1 - \left[1 - e^{-\alpha(\text{arcsech}x)^{\beta}}\right]$$
$$= e^{-\alpha(\text{arcsech}x)^{\beta}}.$$

Note that the hyperbolic secant function is a decreasing function on $(0, \infty)$. The associated pdf follows by differentiating $F(x, \alpha, \beta)$ with respect to x and using $\partial(\text{arcsech}x)/\partial x = -\left(x\sqrt{1-x^2}\right)^{-1}$. Hence, the proof is completed. □

Proof of Proposition 2

Proof. The ASHW distribution is identifiable once $F(x, \alpha_1, \beta_1) = F(x, \alpha_2, \beta_2)$ is valid if and only if $\alpha_1 = \alpha_2$ and $\beta_1 = \beta_2$. After some developments, we get

$$F(x, \alpha_1, \beta_1) = F(x, \alpha_2, \beta_2) \quad \Leftrightarrow \quad e^{-\alpha_1(\text{arcsech}x)^{\beta_1}} = e^{-\alpha_2(\text{arcsech}x)^{\beta_2}}$$
$$\Leftrightarrow \quad \alpha_1 (\text{arcsech}x)^{\beta_1} = \alpha_2 (\text{arcsech}x)^{\beta_2} \quad \Leftrightarrow \quad \frac{\alpha_1}{\alpha_2} (\text{arcsech}x)^{\beta_1 - \beta_2} = 1$$
$$\Leftrightarrow \quad \log\left(\frac{\alpha_1}{\alpha_2}\right) + (\beta_1 - \beta_2)\log(\text{arcsech}\, x) = 0.$$

This equality is satisfied for any x if and only the term varying is x is not present, so $\beta_1 = \beta_2$, and this also implies that $\log(\alpha_1/\alpha_2) = 0$, so $\alpha_1 = \alpha_2$. It is concluded that the model is identifiable. □

Proof of Proposition 3

Proof. Let us investigate the proof of the two items, in turn.

- From Equation (1), we have

$$\frac{\partial}{\partial \alpha} F(x, \alpha, \beta) = F'_{\alpha}(x, \alpha, \beta) = -(\text{arcsech}x)^{\beta} e^{-\alpha(\text{arcsech}x)^{\beta}} < 0.$$

  Hence, $F(x, \alpha, \beta)$ is decreasing with respect to the parameter α. Moreover, we have

$$\frac{\partial^2}{\partial \alpha^2} F(x, \alpha, \beta) = (\text{arcsech}x)^{2\beta} e^{-\alpha(\text{arcsech}x)^{\beta}} > 0,$$

  proving the convexity of $F(x, \alpha, \beta)$ with respect to the parameter α.

- Similarly, we have

$$\frac{\partial}{\partial \beta} F(x, \alpha, \beta) = F'_{\beta}(x, \alpha, \beta) = -\alpha (\text{arcsech}x)^{\beta} \log(\text{arcsech}x) e^{-\alpha(\text{arcsech}x)^{\beta}}.$$

  It is clear that the $\log(\text{arcsech}x)$ term determines the sign of above Equation, and $\log(\text{arcsech}x)$ is positive if $x \in (0, 2e/(e^2+1))$, and negative if $x \in (2e/(e^2+1), 1)$. Hence, it can be concluded that $F(x, \alpha, \beta)$ is decreasing with respect to β if $x \in (0, 2e/(e^2+1))$, and increasing with respect to β if $x \in (2e/(e^2+1), 1)$.

Proposition 3 is proved. □

Proof of Proposition 4

Proof. After simplifications, owing to Equation (2), we arrive at

$$q(x, \alpha_1, \alpha_2, \beta) = \frac{\alpha_1}{\alpha_2} e^{(\alpha_2-\alpha_1)(\text{arcsech}x)^\beta}.$$

Hence

$$\frac{\partial}{\partial x} q(x, \alpha_1, \alpha_2, \beta) = -\beta \frac{\alpha_1}{\alpha_2}(\alpha_2 - \alpha_1)\left(x\sqrt{1-x^2}\right)^{-1} (\text{arcsech}x)^{\beta-1} e^{(\alpha_2-\alpha_1)(\text{arcsech}x)^\beta}.$$

Since $\alpha_2 - \alpha_1 \geq 0$, the above derivative function is negative, implying that $q(x, \alpha_1, \alpha_2, \beta)$ is decreasing. This finishes the proof of Proposition 4. □

Proof of Proposition 5

Proof. The inequality is clear for $x \notin (0, 1)$. For $x \in (0, 1)$, let us notice that

$$\text{arcsech}x = \log\left[\left(1 + \sqrt{1-x^2}\right)/x\right] = -\log x + \log\left(1 + \sqrt{1-x^2}\right) \geq -\log x.$$

Therefore, $\alpha\,(\text{arcsech}x)^\beta \geq \alpha\,(-\log x)^\beta$, which implies that

$$F(x, \alpha, \beta) = e^{-\alpha(\text{arcsech}x)^\beta} \leq e^{-\alpha(-\log x)^\beta} = F_*(x, \alpha, \beta).$$

This terminates the proof of Proposition 5. □

Proof of Proposition 6

Proof. We can write $X = \text{sech}T$ where $T \sim W(\alpha, \beta)$. Therefore, owing to the general binomial theorem with $e^{-2T} \in (0, 1)$ almost surely, we obtain

$$X^r = (\text{sech}T)^r = 2^r \frac{e^{-rT}}{(1 + e^{-2T})^r} = 2^r \sum_{k=0}^{+\infty} \binom{-r}{k} e^{-(r+2k)T}.$$

Hence, by the Fubini-Tonelli theorem, we get

$$m_r = E(X^r) = 2^r \sum_{k=0}^{+\infty} \binom{-r}{k} E(e^{-(r+2k)T}).$$

Now, by the exponential series, Fubini-Tonelli theorem and the definition of the Weibull distribution, we get

$$E(e^{-(r+2k)T}) = \sum_{\ell=0}^{+\infty} \frac{(-1)^\ell}{\ell!}(r + 2k)^\ell E(T^\ell) = \sum_{\ell=0}^{+\infty} \frac{(-1)^\ell}{\ell!}(r + 2k)^\ell \alpha^{-\ell/\beta}\Gamma\left(\frac{\ell}{\beta} + 1\right).$$

The desired result is obtained by putting all the above equalities together. This concludes the proof of Proposition 6. □

Proof of Proposition 7

Proof. By expressing Y_t as $Y_t = \text{sech}\,T$ if $\{\text{sech}\,T \leq t\} = \{T \geq \text{arcsech}\,t\}$ and 0 otherwise, we get

$$m_r(t) = 2^r \sum_{k=0}^{+\infty} \binom{-r}{k} E[e^{-(r+2k)T} I(T \geq \text{arcsech}\,t)],$$

where $I(T \geq \operatorname{arcsech} t) = 1$ if the event $\{T \geq \operatorname{arcsech} t\}$ is realized, and 0 otherwise. We conclude by noticing that

$$
\begin{aligned}
E[e^{-(r+2k)T} I(T \geq \operatorname{arcsech} t)] &= \sum_{\ell=0}^{+\infty} \frac{(-1)^\ell}{\ell!}(r+2k)^\ell E[T^\ell I(T \geq \operatorname{arcsech} t)] \\
&= \sum_{\ell=0}^{+\infty} \frac{(-1)^\ell}{\ell!}(r+2k)^\ell \alpha^{-\ell/\beta} \Gamma\left(\frac{\ell}{\beta}+1, \alpha(\operatorname{arcsech} t)^\beta\right).
\end{aligned}
$$

The proof of Proposition 7 is completed. □

Proof of Proposition 8

Proof. The binomial formula gives

$$
\begin{aligned}
m^*_{s,(j)} &= \int_{-\infty}^{+\infty} x^s f_{X_{(j)}}(x,\alpha,\beta)dx \\
&= c_{i,n} \sum_{k=0}^{n-j} \binom{n-j}{k} (-1)^k \int_0^1 x^s f(x,\alpha,\beta) F(x,\alpha,\beta)^{k+j-1} dx.
\end{aligned}
$$

Now, we can remark

$$
\begin{aligned}
f(x,\alpha,\beta)F(x,\alpha,\beta)^{k+j-1} &= \frac{\alpha\beta}{x\sqrt{1-x^2}} (\operatorname{arcsech} x)^{\beta-1} e^{-\alpha(j+k)(\operatorname{arcsech} x)^\beta} \\
&= \frac{1}{k+j} f(x, \alpha(k+j), \beta).
\end{aligned}
$$

Therefore

$$
m^*_{s,(j)} = c_{i,n} \sum_{k=0}^{n-j} \binom{n-j}{k} (-1)^k \frac{1}{k+j} \int_0^1 x^s f(x,\alpha(k+j),\beta) dx = \sum_{k=0}^{n-j} v_{j,k} m^\dagger_{j,k,s}.
$$

This concludes the proof of Proposition 8. □

Proof of Proposition 9

Proof. By using some basics concepts in probability theory and the expressions of $F(x,\alpha_2,\beta)$ and $f(x,\alpha_1,\beta)$ in Equations (1) and (2), respectively, we get

$$
\begin{aligned}
\tau &= \int_{-\infty}^{+\infty} F(x,\alpha_2,\beta) f(x,\alpha_1,\beta) dx \\
&= \int_0^1 e^{-\alpha_2(\operatorname{arcsech} x)^\beta} \frac{\alpha_1\beta}{x\sqrt{1-x^2}} (\operatorname{arcsech} x)^{\beta-1} e^{-\alpha_1(\operatorname{arcsech} x)^\beta} dx.
\end{aligned}
$$

A rearrangement of the integral with the use of the integral property of the pdf of the $ASHW(\alpha_1+\alpha_2,\beta)$ distribution give

$$
\begin{aligned}
\tau &= \frac{\alpha_1}{\alpha_1+\alpha_2} \int_0^1 \frac{\beta(\alpha_1+\alpha_2)}{x\sqrt{1-x^2}} (\operatorname{arcsech} x)^{\beta-1} e^{-(\alpha_1+\alpha_2)(\operatorname{arcsech} x)^\beta} dx \\
&= \frac{\alpha_1}{\alpha_1+\alpha_2} \int_{-\infty}^{+\infty} f(x,\alpha_1+\alpha_2,\beta) dx = \frac{\alpha_1}{\alpha_1+\alpha_2}.
\end{aligned}
$$

Proposition 9 is proved. □

Score functions, observed information matrix and existence of the MLEs

In theory, MLEs of the α and β parameters follow by solving

$$\frac{\partial \ell(\mathbf{\Lambda})}{\partial \alpha} = \frac{n}{\alpha} - \sum_{i=1}^{n} (\operatorname{arcsech} x_i)^{\beta} = 0$$

and

$$\frac{\partial \ell(\mathbf{\Lambda})}{\partial \beta} = \frac{n}{\beta} + \sum_{i=1}^{n} \operatorname{arcsech} x_i - \alpha \sum_{i=1}^{n} (\operatorname{arcsech} x_i)^{\beta} \log (\operatorname{arcsech} x_i) = 0.$$

From scroe function belong to α parameter , the desired solution satisfied

$$\alpha = \frac{n}{\sum\limits_{i=1}^{n} (\operatorname{arcsech} x_i)^{\beta}}.$$

Then, by combining above equation and log-likelihood function, the following PLL function can be derived:

$$\ell(\beta) = n \log \left(\frac{n}{\sum\limits_{i=1}^{n} (\operatorname{arcsech} x_i)^{\beta}} \right) + n \log \beta - \sum_{i=1}^{n} \log \left[x_i \sqrt{1 - x_i^2} \right] + (\beta - 1) \sum_{i=1}^{n} \operatorname{arcsech} x_i - n.$$

Going on the parameter estimation of the parameter β based on its PLL function, we have

$$\frac{\partial \ell(\beta)}{\partial \beta} = \frac{n}{\beta} - n \frac{\sum\limits_{i=1}^{n} (\operatorname{arcsech} x_i)^{\beta} \log (\operatorname{arcsech} x_i)}{\sum\limits_{i=1}^{n} (\operatorname{arcsech} x_i)^{\beta}} + \sum_{i=1}^{n} \operatorname{arcsech} x_i.$$

Under mild regularity conditions, The MLEs have the bivariate normal distribution with mean $\mu = (\alpha, \beta)$ and covariance matrix I^{-1}, where I denotes the following 2×2 observed information matrix:

$$I = - \left. \begin{pmatrix} \frac{\partial^2}{\partial \alpha^2} \ell(\mathbf{\Lambda}) & \frac{\partial^2}{\partial \alpha \partial \beta} \ell(\mathbf{\Lambda}) \\ \frac{\partial^2}{\partial \alpha \partial \beta} \ell(\mathbf{\Lambda}) & \frac{\partial^2}{\partial \beta^2} \ell(\mathbf{\Lambda}) \end{pmatrix} \right|_{\mathbf{\Lambda} = (\hat{\alpha}, \hat{\beta})},$$

The components of I can be derived through the following derivatives formula:

$$\frac{\partial^2}{\partial \alpha^2} \ell(\mathbf{\Lambda}) = -\frac{n}{\alpha^2},$$

$$\frac{\partial^2}{\partial \alpha \partial \beta} \ell(\mathbf{\Lambda}) = \frac{\partial^2}{\partial \beta \partial \alpha} \ell(\mathbf{\Lambda}) = -\sum_{i=1}^{n} (\operatorname{arcsech} x_i)^{\beta} \log (\operatorname{arcsech} x_i)$$

and

$$\frac{\partial^2}{\partial \beta^2} \ell(\mathbf{\Lambda}) = -\frac{n}{\beta^2} - \alpha \sum_{i=1}^{n} (\operatorname{arcsech} x_i)^{\beta} \log^2 (\operatorname{arcsech} x_i).$$

Now, we discuss the uniqueness and existence of the MLEs. Using the above second derivative of the parameter α, it can be seen clearly that since for all $\alpha > 0$ and n, $\partial^2 \ell(\mathbf{\Lambda}) / \partial \alpha^2 < 0$. This inequality indicates that the $\partial \ell(\mathbf{\Lambda}) / \partial \alpha$ is strictly decreasing in α. Furthermore, $\lim_{\alpha \to 0} \partial \ell(\mathbf{\Lambda}) / \partial \alpha = +\infty$ and $\lim_{\alpha \to +\infty} \partial \ell(\mathbf{\Lambda}) / \partial \alpha = -\sum\limits_{i=1}^{n} (\operatorname{arcsech} x_i)^{\beta} < 0$. Then it is also concluded that $\widehat{\alpha}$ exists and is unique when parameter β is given or known.

On the other hand, using the above second derivative of the parameter β, it can be seen that $\partial^2 \ell(\mathbf{\Lambda}) / \partial \beta^2 < 0$ and this inequality indicates that the $\partial \ell(\mathbf{\Lambda}) / \partial \beta$ is strictly decreasing in β. Furthermore, $\lim_{\beta \to 0} \partial \ell(\mathbf{\Lambda}) / \partial \beta = +\infty$ and $\lim_{\beta \to +\infty} \partial \ell(\mathbf{\Lambda}) \partial \beta = -\infty$. Then it is conluded that $\widehat{\beta}$ exists and is unique when parameter α is given or known.

The competing distributions for univariate data modeling

- Beta distribution:

$$f_{Beta}(x,\alpha,\beta)=\frac{1}{B(\alpha,\beta)}\,(1-x)^{\beta-1}\,x^{\alpha-1},\quad x\in(0,1),$$

  and $f_{Beta}(x,\alpha,\beta)=0$ for $x\notin(0,1)$, where $\alpha>0$, $\beta>0$, and $B(\alpha,\beta)$ is the standard beta function.

- Kw distribution:

$$f_{Kw}(x,\alpha,\beta)=\alpha\beta\,(1-x^{\alpha})^{\beta-1}\,x^{\alpha-1},\quad x\in(0,1),$$

  and $f_{Kw}(x,\alpha,\beta)=0$ for $x\notin(0,1)$, where $\alpha>0$ and $\beta>0$.

- Johnson S_B distribution:

$$f_{S_B}(x,\alpha,\beta)=\frac{\beta}{x\,(1-x)}\phi\left[\beta\log\left(\frac{x}{1-x}\right)+\alpha\right],\quad x\in(0,1),$$

  and $f_{S_B}(x,\alpha,\beta)=0$ for $x\notin(0,1)$, where $\alpha\in\mathbb{R}$, $\beta>0$, and $\phi(x)$ is the pdf of the standard normal distribution.

- UG distribution:

$$f_{UG}(x,\alpha,\beta)=\alpha\beta x^{-\beta-1}e^{-\alpha\left(x^{-\beta}-1\right)},\quad x\in(0,1),$$

  and $f_{UG}(x,\alpha,\beta)=0$ for $x\notin(0,1)$, where $\alpha>0$ and $\beta>0$.

Score vector components of the proposed regression model for MLE method

The derivatives of the Equation (17) with respect to model parameters β and $\boldsymbol{\delta}$ are given by

$$\frac{\partial\ell\left(\boldsymbol{\Delta}\right)}{\partial\beta}=\frac{n}{\beta}-\sum_{i=1}^{n}\log\left(\text{arcsech}\mu_i\right)+\sum_{i=1}^{n}\log\left(\text{arcsech}y_i\right)+\log u\sum_{i=1}^{n}\left(\frac{\text{arcsech}y_i}{\text{arcsech}\mu_i}\right)^{\beta}\log\left(\frac{\text{arcsech}y_i}{\text{arcsech}\mu_i}\right)$$

and

$$\frac{\partial\ell\left(\boldsymbol{\Delta}\right)}{\partial\delta_k}=-\beta\sum_{i=1}^{n}\frac{\partial\text{arcsech}\mu_i/\partial\delta_k}{\text{arcsech}\mu_i}-\beta\log u\sum_{i=1}^{n}\left(\text{arcsech}y_i\right)^{\beta}\left(\text{arcsech}\mu_i\right)^{-\beta-1}\frac{\partial\text{arcsech}\mu_i}{\partial\delta_k},$$

where

$$\begin{aligned}\frac{\partial\text{arcsech}\mu_i}{\partial\delta_k}&=-\left(\frac{e^{\mathbf{x}_i\boldsymbol{\delta}^T}}{1+e^{\mathbf{x}_i\boldsymbol{\delta}^T}}\sqrt{1-\frac{e^{2\mathbf{x}_i\boldsymbol{\delta}^T}}{(1+e^{\mathbf{x}_i\boldsymbol{\delta}^T})^2}}\right)^{-1}\frac{x_{ik}e^{\mathbf{x}_i\boldsymbol{\delta}^T}}{(1+e^{\mathbf{x}_i\boldsymbol{\delta}^T})^2}\\&=-\frac{x_{ik}}{\sqrt{1+2e^{\mathbf{x}_i\boldsymbol{\delta}^T}}}=-x_{ik}\left(\frac{1+\mu_i}{1-\mu_i}\right)^{-1/2}.\end{aligned}$$

Since above Equations consist of the nonlinear function according to model parameters, these log-likelihood functions can be maximized directly by the software such as R and Matlab.

The competing distributions for regression modeling

- The pdf of the beta regression model is given as

$$f_{Beta}(y,\beta,\mu)=\frac{\Gamma\left(\beta\right)}{\Gamma\left(\beta\mu\right)\Gamma\left(\left(1-\mu\right)\beta\right)}y^{\beta\mu-1}\left(1-y\right)^{(1-\mu)\beta-1},\quad y\in(0,1),$$

  $f_{Beta}(y,\beta,\mu)=0$ for $y\notin(0,1)$, where $\mu\in(0,1)$ and $\beta>0$,

- The pdf of the Kw model is specified by

$$f_{Kw}(y,\beta,\mu) = \frac{\beta \log(0.5)}{\log\left(1-\mu^{\beta}\right)} y^{\beta-1} \left(1-y^{\beta}\right)^{\log(0.5)/(\beta(1-\mu)-1)}, \quad y \in (0,1),$$

$f_{Kw}(y,\beta,\mu) = 0$ for $y \notin (0,1)$, where $\mu \in (0,1)$ and $\beta > 0$,

- The pdf of the LEEG model is given as

$$f_{LEEG}(y,\beta,\mu) = \frac{\beta \mu^{\beta} \left(1-\mu^{\beta}\right) y^{\beta-1}}{\left(\mu^{\beta} + \left(1-2\mu^{\beta}\right) y^{\beta}\right)^{2}}, \quad y \in (0,1),$$

$f_{LEEG}(y,\beta,\mu) = 0$ for $y \notin (0,1)$, where $\mu \in (0,1)$ and $\beta > 0$.