Closed form expressions for the continuous ranked probability score

Alexander Jordan

February 13, 2015

The continuous ranked probability score can be given in multiple equivalent forms,

$$\begin{aligned}
\mathrm{CRPS}(F, y) &= \mathbb{E}_F|Y - y| - \frac{1}{2}\mathbb{E}_F|Y - Y'| && (1)\\
&= \int_{-\infty}^{\infty} (F(x) - \mathbb{1}(x \geq y))^2 dx && (2)\\
&= 2\int_0^1 (\mathbb{1}\{y < F^{-1}(\alpha)\} - \alpha)(F^{-1}(\alpha) - y)d\alpha, && (3)
\end{aligned}$$

where (1) is the kernel representation, followed by the threshold decomposition (2), and lastly the quantile decomposition (3). The threshold decomposition corresponds to the integral of the Brier score over all possible thresholds, while the quantile decomposition is the integral of the quantile score over all possible probabilities.

Contents

1 Distributions for discrete variables with infinite support

1.1 Poisson distribution

The probability mass function and cumulative distribution function of the Poisson distribution with parameter $\lambda > 0$, are given by

$$f_\lambda(x) = \begin{cases} \frac{\lambda^x}{x!} e^{-\lambda} & x = 0, 1, 2, ... \\ 0 & \text{else} \end{cases}$$

$$F_\lambda(x) = \begin{cases} \frac{\Gamma_u(\lfloor x+1 \rfloor, \lambda)}{\Gamma(\lfloor x+1 \rfloor)} & x \geq 0 \\ 0 & x < 0 \end{cases},$$

where Γ denotes the gamma function, Γ_u denotes the upper incomplete gamma function, and $\lfloor \cdot \rfloor$ denotes the floor function.

$$\text{CRPS}(F_\lambda, y) = (y - \lambda)\Big(2F_\lambda(y) - 1\Big) + \lambda \left[2f_\lambda(\lfloor y \rfloor) - e^{-2\lambda}\Big(I_0(2\lambda) + I_1(2\lambda)\Big)\right]$$

where I_m denotes the modified Bessel function of the first kind

The closed form expression for the CRPS is given by Wei and Held (2014).

1.2 Negative binomial distribution

The probability mass function and cumulative distribution function of the negative binomial distribution with number of successes $n > 0$ and success probability $p \in (0, 1]$, where the distribution describes the number of failures until the target number of successes is reached, are given by

$$f_{n,p}(x) = \begin{cases} \frac{\Gamma(x+n)}{\Gamma(n)x!} p^n (1-p)^x & x = 0, 1, 2, ... \\ 0 & \text{else} \end{cases}$$

$$F_{n,p}(x) = \begin{cases} I_p(n, \lfloor x+1 \rfloor) & x \geq 0 \\ 0 & x < 0 \end{cases},$$

where I_x denotes the regularized incomplete beta function, and $\lfloor \cdot \rfloor$ denotes the floor function.

$\mathrm{CRPS}(F_{n,p}, y) =$

$$y\Big(2F_{n,p}(y) - 1\Big) - \tfrac{n(1-p)}{p^2}\Big[p\Big(2F_{n+1,p}(y-1) - 1\Big) + {}_2F_1\left(n+1, \tfrac{1}{2}; 2; -\tfrac{4(1-p)}{p^2}\right)\Big]$$

where ${}_2F_1$ denotes the hypergeometric function

The closed form expression for the CRPS is given by Wei and Held (2014).

2 Distributions for variables on a bounded interval

2.1 Continuous uniform

The cumulative distribution function of the continuous uniform distribution with lower bound $a \in \mathbb{R}$ and upper bound $b \in \mathbb{R}$, is given by

$$F_{a,b}(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \leq x < b \\ 1 & x \geq b \end{cases}$$

and the CRPS can be calculated using any of the three representations.

$$\mathrm{CRPS}(F_{a,b},y) = (y-a)\Big(2F_{a,b}(y)-1\Big) + (b-a)\Big(1/3 - F_{a,b}(y)^2\Big)$$

2.2 Beta distribution

The cumulative distribution function of the beta distribution with parameters $\alpha, \beta > 0$ is given by

$$F_{\alpha,\beta}(x) = \begin{cases} 0 & x < 0 \\ I_x(\alpha,\beta) & 0 \leq x < 1 \\ 1 & x \geq 1 \end{cases},$$

where I_x denotes the regularized incomplete beta function.

$$\mathrm{CRPS}(F_{\alpha,\beta},y) = y(2F_{\alpha,\beta}(y)-1) + \tfrac{\alpha}{\alpha+\beta}\Big(1 - 2F_{\alpha+1,\beta}(y) - \tfrac{2B(2\alpha,2\beta)}{\alpha B(\alpha,\beta)^2}\Big)$$

where B denotes the beta function

3 Distributions for variables on the real line

3.1 Laplace distribution

The cdf of the Laplace distribution with location parameter $\mu \in \mathbb{R}$ and scale parameter $\sigma > 0$ is given by

$$F_{\mu,\sigma}(x) = \begin{cases} \frac{1}{2}\exp\left(\frac{x-\mu}{\sigma}\right) & x < \mu \\ 1 - \frac{1}{2}\exp\left(-\frac{x-\mu}{\sigma}\right) & x \geq \mu. \end{cases}$$

$$\mathrm{CRPS}(F_{\mu,\sigma},y) = \sigma\Big[\tfrac{y-\mu}{\sigma}\Big(2F_{\mu,\sigma}(y)-1\Big) - 2p_{\min}\Big(\log(2p_{\min})-1\Big) - \tfrac{3}{4}\Big]$$

where $p_{\min} = \min\Big(F_{\mu,\sigma}(y), 1 - F_{\mu,\sigma}(y)\Big)$

The CRPS can be calculated using the quantile decomposition.

3.2 Logistic distribution

The cdf of the Logistic distribution with location parameter $\mu \in \mathbb{R}$ and scale parameter $\sigma > 0$ is given by

$$F_{\mu,\sigma}(x) = \frac{1}{1 + \exp(-\frac{x-\mu}{\sigma})}.$$

$$\mathrm{CRPS}(F_{\mu,\sigma}, y) = \sigma\Big[\tfrac{y-\mu}{\sigma}\Big(2p_y - 1\Big) - 1 - 2\Big(p_y \log(p_y) + (1-p_y)\log(1-p_y)\Big)\Big]$$

where $p_y = F_{\mu,\sigma}(y)$

The CRPS can be calculated using the quantile decomposition.

3.3 Normal distribution

As usual, φ denotes the probability density function and Φ denotes the cumulative distribution function of the standard normal distribution, which can be transformed by parameters $\mu \in \mathbb{R}$ and $\sigma > 0$ to change mean and variance.

$$\mathrm{CRPS}(\mathcal{N}_{\mu,\sigma}, y) = \sigma\Big[\tfrac{y-\mu}{\sigma}\Big(2\Phi\Big(\tfrac{y-\mu}{\sigma}\Big) - 1\Big) + 2\varphi\Big(\tfrac{y-\mu}{\sigma}\Big) - \tfrac{1}{\sqrt{\pi}}\Big]$$

Gneiting et al. (2005) showed that the CRPS can be calculated using the kernel representation by computing

$$\mathbb{E}_{\mathcal{N}_{\mu,\sigma}}|Y - y| = \sigma\Big[\tfrac{y-\mu}{\sigma}\Big(2\Phi\left(\tfrac{y-\mu}{\sigma}\right) - 1\Big) + 2\varphi\Big(\tfrac{y-\mu}{\sigma}\Big)\Big]$$

and

$$\mathbb{E}_{\mathcal{N}_{\mu,\sigma}}|Y - Y'| = \sigma\frac{2}{\sqrt{\pi}}.$$

3.3.1 Finite mixture of normal distributions

Grimit et al. (2006) give an expression for the CRPS when the predictive cumulative distribution function is a mixture of M normal distributions with weights $\omega_1, ..., \omega_M$, $\sum_{i=1}^M \omega_i = 1$, i.e.

$$F(x) = \sum_{i=1}^{M} \omega_i \Phi\left(\frac{x-\mu_i}{\sigma_i}\right).$$

$$\mathrm{CRPS}(F, y) = \sum_{i=1}^{M} \omega_i A\Big(y - \mu_i, \sigma_i^2\Big) - \frac{1}{2}\sum_{i=1}^{M}\sum_{j=1}^{M} \omega_i \omega_j A\Big(\mu_i - \mu_j, \sigma_i^2 + \sigma_j^2\Big)$$

where $A\Big(\mu, \sigma^2\Big) = \sigma\Big[\frac{\mu}{\sigma}\Big(2\Phi\Big(\frac{\mu}{\sigma}\Big) - 1\Big) + 2\varphi\Big(\frac{\mu}{\sigma}\Big)\Big] = \mathbb{E}_{\mathcal{N}_{\mu,\sigma}}|Y|$

Since the normal distribution is a stable distribution the CRPS can be expressed in terms of a function $A(\mu, \sigma^2)$, which is simply the expectation of the absolute value of a normal random variable with mean μ and variance σ^2. Note that the computation time increases quadratic in M, so that numerical integration of the threshold decomposition (only linear computational complexity in M) to machine precision may be faster than using the formula above when the mixture consists of more than several thousand components.

3.3.2 Split-normal distribution

The cumulative distribution function of the split-normal, or two-piece normal distribution, with location parameter $\mu \in \mathbb{R}$ and scale parameters $\sigma_1, \sigma_2 > 0$ is given by

$$F_{\mu,\sigma_1,\sigma_2}(x) = \begin{cases} \frac{2\sigma_1}{\sigma_1+\sigma_2}\Phi\left(\frac{x-\mu}{\sigma_1}\right) & \text{if } x < \mu, \\ \frac{\sigma_1-\sigma_2}{\sigma_1+\sigma_2} + \frac{2\sigma_2}{\sigma_1+\sigma_2}\Phi\left(\frac{x-\mu}{\sigma_2}\right) & \text{if } x \geq \mu. \end{cases}$$

$$
\mathrm{CRPS}(F_{\mu,\sigma_1,\sigma_2}, y) = \begin{cases} \frac{4\sigma_1^2}{\sigma_1+\sigma_2}\left[\frac{y-\mu}{\sigma_1}\Phi\left(\frac{y-\mu}{\sigma_1}\right) + \varphi\left(\frac{y-\mu}{\sigma_1}\right)\right] \\ \quad -(y-\mu) \\ \quad + \frac{2}{\sqrt{\pi}}\frac{\sqrt{2}\sigma_2(\sigma_2^2-\sigma_1^2)-(\sigma_1^3+\sigma_2^3)}{(\sigma_1+\sigma_2)^2} & \text{if} \quad y \le \mu \\ \frac{4\sigma_2^2}{\sigma_1+\sigma_2}\left[\frac{y-\mu}{\sigma_2}\Phi\left(\frac{y-\mu}{\sigma_2}\right) + \varphi\left(\frac{y-\mu}{\sigma_2}\right)\right] \\ \quad +(y-\mu)\frac{(\sigma_1-\sigma_2)^2-4\sigma_2^2}{(\sigma_1+\sigma_2)^2} \\ \quad + \frac{2}{\sqrt{\pi}}\frac{\sqrt{2}\sigma_1(\sigma_1^2-\sigma_2^2)-(\sigma_1^3+\sigma_2^3)}{(\sigma_1+\sigma_2)^2} & \text{if} \quad y \ge \mu \end{cases}
$$

The CRPS can be calculated using the kernel representation (Gneiting and Thorarinsdottir, 2010) by computing

$$
\begin{aligned}
&\mathbb{E}_{F_{\mu,\sigma_1,\sigma_2}}|Y-y| \\
&\quad = \frac{2}{\sigma_1+\sigma_2}\left[\frac{\sigma_1^2+\sigma_2^2}{\sqrt{2\pi}} + \frac{(y-\mu)(\sigma_1-\sigma_2)}{2} + h\Big(\mathbb{1}_{\{\mu>y\}}\sigma_1 + \mathbb{1}_{\{\mu<y\}}\sigma_2\Big)\right],
\end{aligned}
$$

where

$$
h(\sigma) = \sigma^2\left(2\left[\varphi\left(\frac{y-\mu}{\sigma}\right) - \frac{1}{\sqrt{2\pi}}\right] + \frac{y-\mu}{\sigma}\left[2\Phi\left(\frac{y-\mu}{\sigma}\right) - 1\right]\right),
$$

and

$$
\mathbb{E}_{F_{\mu,\sigma_1,\sigma_2}}|Y-Y'| = \left(\frac{(2-\sqrt{2})(\sigma_1^3+\sigma_2^3) + \sqrt{2}(\sigma_1^2\sigma_2+\sigma_1\sigma_2^2)}{(\sigma_1+\sigma_2)^2}\right)\frac{2}{\sqrt{\pi}}.
$$

3.4 Student's t-distribution

The pdf and cdf of a variable Z with non-standardized t-distribution, i.e. such that $\frac{Z-\mu}{\sigma}$ follows a standard t-distribution, are given by

$$
\begin{aligned}
f_{\nu,\mu,\sigma}(z) &= \frac{1}{\sqrt{\nu\sigma^2}B(\frac{1}{2},\frac{\nu}{2})}\left(1+\frac{(z-\mu)^2}{\nu\sigma^2}\right)^{-\frac{\nu+1}{2}} \\
F_{\nu,\mu,\sigma}(z) &= \frac{1}{2} + \frac{(z-\mu)\ {}_2F_1(\frac{1}{2},\frac{\nu+1}{2};\frac{3}{2};-\frac{(z-\mu)^2}{\nu\sigma^2})}{\sqrt{\nu\sigma^2}B(\frac{1}{2},\frac{\nu}{2})},
\end{aligned}
$$

where $\nu > 0$ denotes the degrees of freedom, $\mu \in \mathbb{R}$ denotes the mean, and $\sigma > 0$ is a scaling parameter. The symbol B denotes the beta function and ${}_2F_1$ denotes the hypergeometric function.

$$\text{CRPS}(F_{\nu,\mu,\sigma}, y) =$$
$$\sigma\left[2f_\nu\left(\frac{y-\mu}{\sigma}\right)\frac{\nu+(\frac{y-\mu}{\sigma})^2}{\nu-1}+\frac{y-\mu}{\sigma}\left(2F_\nu\left(\frac{y-\mu}{\sigma}\right)-1\right)-\frac{2\sqrt{\nu}}{\nu-1}\frac{B(\frac{1}{2},\nu-\frac{1}{2})}{B(\frac{1}{2},\frac{\nu}{2})^2}\right]$$

Based on the kernel representation, we can compute the CRPS for such a non-standardized t-distribution with $\nu > 1$, by using the identities

$$\mathbb{E}_{F_{\nu,\mu,\sigma}}|Y-y| = \sigma\left[2f_\nu\left(\frac{y-\mu}{\sigma}\right)\frac{\nu+(\frac{y-\mu}{\sigma})^2}{\nu-1}+\frac{y-\mu}{\sigma}\left(2F_\nu\left(\frac{y-\mu}{\sigma}\right)-1\right)\right] \tag{4}$$

and

$$\mathbb{E}_{F_{\nu,\mu,\sigma}}|Y-Y'| = \sigma\frac{4\sqrt{\nu}}{\nu-1}\frac{B(\frac{1}{2},\nu-\frac{1}{2})}{B(\frac{1}{2},\frac{\nu}{2})^2}, \tag{5}$$

where f_ν denotes the pdf of the standard t-distribution, and F_ν denotes its cdf.

3.4.1 Finite mixture of t-distributions (incomplete)

Suppose that the predictive distribution F takes the form of a mixture of $n \geq 2$ pdfs of non-standardized t-distributions with equal degrees of freedom, i.e.

$$F = \sum_{i=1}^{n} \omega_i F_i$$

where $\omega_i \in (0,1)$ with $\sum_{i=1}^n \omega_i = 1$, and $F_i(z) = F_\nu(\frac{z-\mu_i}{\sigma_i})$. In this case, the first part of the kernel representation of the CRPS can be calculated as

$$\mathbb{E}_F|Y-y| = \sum_{i=1}^{n} \omega_i \; \mathbb{E}_{F_i}|Y-y|,$$

inserting the results for the non-standardized t-distribution. Similarly, the second part can be calculated as

$$\begin{aligned}\mathbb{E}_F|Y-Y'| &= \sum_{i=1}^{n}\sum_{j=1}^{n} \omega_i \; \omega_j \; \mathbb{E}_{F_\nu}|\sigma_i Y + \mu_i - \sigma_j Y' - \mu_j| \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} \omega_i \; \omega_j \; \mathbb{E}_{F^*_{ij}}|(\sigma_i+\sigma_j)Y + \mu_{ij}|,\end{aligned}$$

where $\mu_{ij} = \mu_i - \mu_j$, and F^*_{ij} is the pdf of the random variable $(\frac{\sigma_i}{\sigma_i+\sigma_j}X + \frac{\sigma_j}{\sigma_i+\sigma_j}X')$ with independent draws X, X' from a standard t-distribution with ν degrees of freedom. This pdf is given in Ruben (1960) as

$$F^*_{ij}(z) = \int_0^1 \frac{(t(1-t))^{\nu/2-1}}{B(\frac{\nu}{2},\frac{\nu}{2})} F_{2\nu}(g_{ij}(t)z)\ dt$$

where

$$g_{ij}(t) = (\sigma_i + \sigma_j)\sqrt{\frac{2t(1-t)}{\sigma_i^2(1-t)+\sigma_j^2 t}}.$$

Since this distribution is symmetric about 0, we can write

$$\mathbb{E}_{F^*_{ij}}|(\sigma_i+\sigma_j)Y + \mu_{ij}| = \mu_{ij}\left(2F^*_{ij}\left(\tfrac{\mu_{ij}}{\sigma_i+\sigma_j}\right) - 1\right) - 2\int_{-\infty}^{\mu_{ij}} x\ dF^*_{ij}\left(\tfrac{x}{\sigma_i+\sigma_j}\right),$$

leading to

$$\begin{aligned}\mu_{ij}\left(2F^*_{ij}\left(\tfrac{\mu_{ij}}{\sigma_i+\sigma_j}\right) - 1\right) =& \frac{2\mu_{ij}^2}{B(\frac{v}{2},\frac{v}{2})B(\frac{1}{2},\frac{v}{2})\sqrt{\nu}} \\ &\times \int_0^1 \frac{(t(1-t))^{(\nu-1)/2}}{\sqrt{\sigma_i^2(1-t)+\sigma_j^2 t}}\ {}_2F_1\left(\tfrac{1}{2}, \nu+\tfrac{1}{2}; \tfrac{3}{2}; -\frac{\mu_{ij}^2 t(1-t)}{\nu(\sigma_i^2(1-t)+\sigma_j^2 t)}\right)\ dt\end{aligned}$$

and

$$\begin{aligned}\int_{-\infty}^{\mu_{ij}} x\ dF^*_{ij}\left(\tfrac{x}{\sigma_i+\sigma_j}\right) =& -\frac{\sqrt{2\nu}}{B(\frac{\nu}{2},\frac{\nu}{2})B(\frac{1}{2},\frac{\nu}{2})(2\nu-1)} \\ &\times \int_0^1 (t(1-t))^{\nu/2-1}\left(1 + \frac{\mu_{ij}^2 t(1-t)}{\nu(\sigma_i^2(1-t)+\sigma_j^2 t)}\right)^{-\frac{2\nu-1}{2}}\ dt\end{aligned}$$

4 Distributions for non-negative variables

4.1 Exponential distribution

The cumulative distribution function of the exponential distribution with parameter $\lambda > 0$ is given by

$$F_\lambda(x) = \begin{cases} 1 - e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}.$$

$$\text{CRPS}(F_\lambda, y) = |y| + \tfrac{1}{2\lambda}\Big(1 - 4F_\lambda(y)\Big)$$

For calculation of the CRPS using the kernel representation, we compute the identities

$$\mathbb{E}_{F_\lambda}|Y - y| = |y| + \tfrac{1}{\lambda}\Big(1 - 2F_\lambda(y)\Big)$$

and

$$\mathbb{E}_{F_\lambda}|Y - Y'| = \tfrac{1}{\lambda},$$

or the quantile decomposition. For a location-shifted exponential distribution, we refer to the generalized Pareto distribution with shape parameter 0.

4.2 Gamma distribution

The cumulative distribution function of the gamma distribution with shape parameter $\alpha > 0$ and rate parameter $\beta > 0$ is given by

$$F_{\alpha,\beta}(x) = \begin{cases} \frac{\Gamma_l(\alpha,\beta x)}{\Gamma(\alpha)} & x \geq 0 \\ 0 & x < 0 \end{cases},$$

where Γ denotes the gamma function, and Γ_l denotes the lower incomplete gamma function.

$$\text{CRPS}(F_{\alpha,\beta}, y) = y\Big(2F_{\alpha,\beta}(y) - 1\Big) - \tfrac{1}{\beta}\Big[\alpha\Big(2F_{\alpha+1,\beta}(y) - 1\Big) - B(\tfrac{1}{2},\alpha)^{-1}\Big]$$

where B denotes the beta function

The CRPS can be calculated using the kernel representation (Möller and Scheuerer, 2013), where the two components are given by

$$\mathbb{E}_{F_{\alpha,\beta}}|Y - y| = y(2F_{\alpha,\beta}(y) - 1) - \frac{\alpha}{\beta}(2F_{\alpha+1,\beta}(y) - 1)$$

and

$$\mathbb{E}_{F_{\alpha,\beta}}|Y - Y'| = \frac{2}{\beta B(\frac{1}{2},\alpha)}.$$

4.3 Log-distributions

4.3.1 Log-Laplace distribution

The cumulative distribution function of the Log-Laplace distribution with parameters $\mu \in \mathbb{R}$ and $\sigma > 0$ is given by

$$F_{\alpha,\beta}(x) = \begin{cases} 0 & x \leq 0 \\ \frac{1}{2}\exp\left(\frac{\log x - \mu}{\sigma}\right) & 0 < x < e^{\mu} \\ 1 - \frac{1}{2}\exp\left(-\frac{\log x - \mu}{\sigma}\right) & x \geq e^{\mu} \end{cases}.$$

$$\mathrm{CRPS}(F_{\mu,\sigma}, y) = y\Big(2F_{\mu,\sigma}(y) - 1\Big) + e^{\mu}\Big(\frac{\sigma}{4-\sigma^2} + A(y)\Big)$$

where

$$A(y) = \begin{cases} \frac{1}{1+\sigma}\Big(1 - \big[2F_{\mu,\sigma}(y)\big]^{1+\sigma}\Big) & y < \alpha \\ -\frac{1}{1-\sigma}\Big(1 - \big[2(1 - F_{\mu,\sigma}(y))\big]^{1-\sigma}\Big) & y \geq \alpha \end{cases}$$

The CRPS exists for $\sigma < 1$, satisfying the finite-first-moment condition, and the formula can be found using the quantile decomposition.

4.3.2 Log-logistic distribution

The cdf of the Log-logistic distribution with parameters $\mu \in \mathbb{R}$ and $\sigma > 0$ is given by

$$F_{\mu,\sigma}(x) = \begin{cases} \frac{1}{1+\exp\left(-\frac{\log x - \mu}{\sigma}\right)} & x > 0 \\ 0 & x \leq 0 \end{cases}.$$

$$\mathrm{CRPS}(F_{\mu,\sigma}, y) = y\Big(2F_{\mu,\sigma}(y) - 1\Big) - 2e^{\mu}B(1+\sigma, 1-\sigma)\Big(\frac{\sigma-1}{2} + I_{F_{\mu,\sigma}(y)}(1+\sigma, 1-\sigma)\Big)$$

where B denotes the beta function, and I_x denotes the regularized incomplete beta function

The CRPS exists for $\sigma < 1$, satisfying the finite-first-moment condition, and the formula can be found using the quantile decomposition.

4.3.3 Log-normal distribution

The cumulative distribution function of the log-normal distribution with parameters $\mu \in \mathbb{R}$ and $\sigma > 0$ is given in terms of the of the cumulative distribution function Φ of the standard normal distribution,

$$F_{\mu,\sigma}(x) = \begin{cases} \Phi\left(\frac{\log x - \mu}{\sigma}\right) & x > 0 \\ 0 & x \leq 0. \end{cases}$$

$$\boxed{\mathrm{CRPS}(F_{\mu,\sigma}, y) = y\Big(2F_{\mu,\sigma}(y) - 1\Big) - 2\mathrm{e}^{\mu+\sigma^2/2}\left[\Phi\left(\frac{\log y - \mu - \sigma^2}{\sigma}\right) + \Phi\left(\frac{\sigma}{\sqrt{2}}\right) - 1\right]}$$

The closed for expression for the CRPS is given by Baran and Lerch (2014).

4.4 Truncated normal distribution

The truncated normal distribution with location parameter $\mu \in \mathbb{R}$ and scale parameter $\sigma > 0$ takes the mass of the negative tail of the corresponding normal distribution and redistributes it to the positive tail. Formally, the cumulative distribution function is given by

$$F_{\mu,\sigma}(x) = \begin{cases} \frac{\Phi\left(\frac{x-\mu}{\sigma}\right) - \Phi\left(\frac{-\mu}{\sigma}\right)}{\Phi\left(\frac{\mu}{\sigma}\right)} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

$$\boxed{\begin{aligned}&\mathrm{CRPS}(F_{\mu,\sigma}, y) = \\ &\qquad \sigma\left[\Phi\left(\frac{\mu}{\sigma}\right)\right]^{-2}\left[\frac{y-\mu}{\sigma}\Phi\left(\frac{\mu}{\sigma}\right)\left[2\Phi\left(\frac{y-\mu}{\sigma}\right) + \Phi\left(\frac{\mu}{\sigma}\right) - 2\right] + 2\varphi\left(\frac{y-\mu}{\sigma}\right)\Phi\left(\frac{\mu}{\sigma}\right) - \frac{1}{\sqrt{\pi}}\Phi\left(\frac{\sqrt{2}\mu}{\sigma}\right)\right]\end{aligned}}$$

The closed for expression of the CRPS is given by Thorarinsdottir and Gneiting (2010).

4.5 Left-censored normal distribution

Similarly to the truncated normal distribution, the left-censored normal distribution with location parameter $\mu \in \mathbb{R}$ and scale parameter $\sigma > 0$ takes the mass of the negative tail of the corresponding normal distribution and places it as a point-mass at 0. Formally, the cumulative distribution function is given by

$$F_{\mu,\sigma}(x) = \begin{cases} \Phi\left(\frac{x-\mu}{\sigma}\right) & x \geq 0 \\ 0 & x \leq 0 \end{cases}$$

$$\boxed{\mathrm{CRPS}(F_{\mu,\sigma}, y) = \mathrm{CRPS}(\mathcal{N}_{\mu,\sigma}, y) - 2\sigma\varphi\Big(\frac{\mu}{\sigma}\Big)\Phi\Big(-\frac{\mu}{\sigma}\Big) + \frac{\sigma}{\sqrt{\pi}}\Phi\Big(-\sqrt{2}\frac{\mu}{\sigma}\Big) + \mu\Big[\Phi\Big(\frac{\mu}{\sigma}\Big)\Big]^2}$$

The closed for expression of the CRPS is given by Gneiting et al. (2006).

5 Distributions with variable support

5.1 Generalized pareto distribution (GPD)

when $\xi = 0$:

$$\mathrm{CRPS}(F_{\mu,\sigma,0}, y) = |y - \mu| - \sigma \left(2F_{\mu,\sigma,0}(y) - \frac{1}{2}\right)$$

when $\xi \neq 0$:

$$\begin{aligned}\mathrm{CRPS}(F_{\mu,\sigma,\xi}, y) = &\left(y - \mu + \frac{\sigma}{\xi}\right)(2F_{\mu,\sigma,\xi}(y) - 1) \\ &- \frac{2\sigma}{\xi(\xi-1)}\left(\frac{1}{\xi-2} + \left(1 - F_{\mu,\sigma,\xi}(y)\right)^{1-\xi}\right)\end{aligned}$$

The cdf of the generalized pareto distribution with location parameter $\mu \in \mathbb{R}$ and scale parameter $\sigma > 0$ is given dependent on the shape parameter $\xi \in \mathbb{R}$

$$\text{for } \xi = 0: \quad F_{\mu,\sigma,0}(x) = \begin{cases} 0 & x < \mu \\ 1 - \exp(-\frac{x-\mu}{\sigma}) & x \geq \mu \end{cases}$$

$$\text{for } \xi > 0: \quad F_{\mu,\sigma,\xi}(x) = \begin{cases} 0 & x < \mu \\ 1 - (1 + \xi\frac{x-\mu}{\sigma})^{-1/\xi} & x \geq \mu \end{cases}$$

$$\text{for } \xi < 0: \quad F_{\mu,\sigma,\xi}(x) = \begin{cases} 0 & x < \mu \\ 1 - (1 + \xi\frac{x-\mu}{\sigma})^{-1/\xi} & x \in [\mu, \mu - \frac{\sigma}{\xi}] \\ 1 & x > \mu - \frac{\sigma}{\xi} \end{cases}$$

The CRPS can be calculated for $\xi < 1$, satisfying the finite-first-moment condition, and Friederichs and Thorarinsdottir (2012) compute the closed form expressions of the CRPS using the quantile representation. The formulas above are slightly modified to give the correct CRPS values for observations $y < \mu$.

5.2 Generalized extreme value distribution (GEV)

when $\xi = 0$:

$$\mathrm{CRPS}(F_{\mu,\sigma,0}, y) = \mu - y + \sigma(C - \log 2) - 2\sigma Ei(\log F_{\mu,\sigma,0}(y))$$

when $\xi \neq 0$:

$$\begin{aligned}\mathrm{CRPS}(F_{\mu,\sigma,\xi}, y) = &\left(\mu - y - \frac{\sigma}{\xi}\right)(1 - 2F_{\mu,\sigma,\xi}(y)) \\ &- \frac{\sigma}{\xi}\left(2^{\xi}\Gamma(1-\xi) - 2\Gamma_l(1-\xi, -\log F_{\mu,\sigma,\xi}(y))\right)\end{aligned}$$

The cdf of the generalized extreme value distribution with location parameter $\mu \in \mathbb{R}$ and scale parameter $\sigma > 0$ is given dependent on the shape parameter $\xi \in \mathbb{R}$

$$\begin{aligned}
&\text{for } \xi = 0: \quad F_{\mu,\sigma,\xi}(x) = \exp\left(-\exp(-\tfrac{x-\mu}{\sigma})\right) \\
&\text{for } \xi > 0: \quad F_{\mu,\sigma,\xi}(x) = \begin{cases} 0 & x \leq \mu - \frac{\sigma}{\xi} \\ \exp\left(-(1 + \xi\frac{x-\mu}{\sigma})^{-1/\xi}\right) & x > \mu - \frac{\sigma}{\xi} \end{cases} \\
&\text{for } \xi < 0: \quad F_{\mu,\sigma,\xi}(x) = \begin{cases} \exp\left(-(1 + \xi\frac{x-\mu}{\sigma})^{-1/\xi}\right) & x < \mu - \frac{\sigma}{\xi} \\ 1 & x \geq \mu - \frac{\sigma}{\xi}. \end{cases}
\end{aligned}$$

The CRPS can be calculated for $\xi < 1$, satisfying the finite-first-moment condition, and Friederichs and Thorarinsdottir (2012) compute the closed form expressions of the CRPS using the quantile representation, where C is the Euler-Mascheroni constant, $Ei(x) = \int_{-\infty}^{x} \frac{e^t}{t} dt$ is the exponential integral, Γ denotes the gamma function, and Γ_l denotes the lower incomplete gamma function.

5.2.1 Left-censored GEV distribution

when $\xi \neq 0$:

$$\mathrm{CRPS}(F_{\mu,\sigma,\xi}, y) = (\mu - y)(1 - 2p_y) + \mu p_0^2 - 2\frac{\sigma}{\xi}\Big(1 - p_y - \Gamma_l(1-\xi, -\log p_y)\Big) + \frac{\sigma}{\xi}\Big(1 - p_0^2 - 2^{\xi}\Gamma_l(1-\xi, -2\log p_0)\Big)$$

where $p_0 = F_{\mu,\sigma,\xi}(0)$ and $p_y = F_{\mu,\sigma,\xi}(y)$

$$F_{\mu,\sigma,\xi}(x) = \begin{cases} F^{\mathrm{GEV}}_{\mu,\sigma,\xi}(x) & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Scheuerer (2014)

References

Baran, S. and Lerch, S. (2014). Log-normal distribution based emos models for probabilistic wind speed forecasting. *arXiv preprint arXiv:1407.3252*.

Friederichs, P. and Thorarinsdottir, T. L. (2012). Forecast verification for extreme value distributions with an application to probabilistic peak wind prediction. *Environmetrics*, 23:579–594.

Gneiting, T., Larson, K., Westrick, K., Genton, M. G., and Aldrich, E. (2006). Calibrated probabilistic forecasting at the stateline wind energy center: The regime-switching space–time method. *Journal of the American Statistical Association*, 101:968–979.

Gneiting, T., Raftery, A. E., Westveld III, A. H., and Goldman, T. (2005). Calibrated probabilistic forecasting using ensemble model output statistics and minimum crps estimation. *Monthly Weather Review*, 133:1098–1118.

Gneiting, T. and Thorarinsdottir, T. L. (2010). Predicting inflation: Professional experts versus no-change forecasts. *arXiv preprint arXiv:1010.2318*.

Gradshteyn and Ryzhik (2007). *Table of Integrals, Series, and Products*.

Grimit, E. P., Gneiting, T., Berrocal, V., and Johnson, N. A. (2006). The continuous ranked probability score for circular variables and its application to mesoscale forecast ensemble verification. *Quarterly Journal of the Royal Meteorological Society*, 132:2925–2942.

Milgram, M. (2010). On hypergeometrics 3f2 (1)-a review. *arXiv preprint arXiv:1011.4546*.

Möller, D. and Scheuerer, M. (2013). Postprocessing of ensemble forecasts for wind speed over germany.

Ruben, H. (1960). On the distribution of the weighted difference of two independent student variables. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 188–194.

Scheuerer, M. (2014). Probabilistic quantitative precipitation forecasting using ensemble model output statistics. *Quarterly Journal of the Royal Meteorological Society*, 140:1086–1096.

Thorarinsdottir, T. L. and Gneiting, T. (2010). Probabilistic forecasts of wind speed: ensemble model output statistics by using heteroscedastic censored regression. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 173:371–388.

Wei, W. and Held, L. (2014). Calibration tests for count data. *TEST*, 23:787–805.

A Proofs

A.1 Beta distribution

Proof

$$\begin{aligned}
\mathbb{E}_{F_{\alpha,\beta}}|Y-y| &= \int_y^\infty (x-y)f_{\alpha,\beta}(x)dx - \int_{-\infty}^y (x-y)f_{\alpha,\beta}(x)dx \\
&= y(2F_{\alpha,\beta}(y)-1) + \int_y^\infty xf_{\alpha,\beta}(x)dx - \int_{-\infty}^y xf_{\alpha,\beta}(x)dx
\end{aligned}$$

With $y' = \max(\min(y,1),0)$ we get

$$\begin{aligned}
&\int_{y'}^1 xf_{\alpha,\beta}(x)dx - \int_0^{y'} xf_{\alpha,\beta}(x)dx \\
&= \tfrac{B(\alpha+1,\beta)}{B(\alpha,\beta)} \left(\Big(I_1(\alpha+1,\beta) - I_{y'}(\alpha+1,\beta) \Big) - I_{y'}(\alpha+1,\beta) \right) \\
&= \tfrac{\alpha}{\alpha+\beta} \Big(1 - 2F_{\alpha+1,\beta}(y) \Big)
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}_{F_{\alpha,\beta}}|Y-Y'| &= 2\int_0^1 \int_{-\infty}^\infty s' f_{\alpha,\beta}(t) f_{\alpha,\beta}(t-s')dtds' \\
&= 2\int_{-\infty}^\infty f_{\alpha,\beta}(t) \int_{t-1}^t (t-s) f_{\alpha,\beta}(s)dsdt \\
&= 2\int_0^1 f_{\alpha,\beta}(t) \int_0^t (t-s) f_{\alpha,\beta}(s)dsdt \\
&= 2\left(\int_0^1 tf_{\alpha,\beta}(t)F_{\alpha,\beta}(t)dt - \int_0^1 f_{\alpha,\beta}(t) \tfrac{B_t(\alpha+1,\beta)}{B(\alpha,\beta)} dt \right) \\
&= 2\left(\Big[\tfrac{B_t(\alpha+1,\beta)}{B(\alpha,\beta)} F_{\alpha,\beta}(t) \Big]_0^1 - 2\int_0^1 f_{\alpha,\beta}(t) \tfrac{B_t(\alpha+1,\beta)}{B(\alpha,\beta} dt \right) \\
&= 2\left(\tfrac{\alpha}{\alpha+\beta} - \tfrac{2}{B(\alpha,\beta)^2} \int_0^1 t^{\alpha-1}(1-t)^{\beta-1} B_t(\alpha+1,\beta)dt \right)
\end{aligned}$$

Using equation 8.17.8 from the Digital Library of Mathematical Functions, and equation 7.512.5 in Gradshteyn and Ryzhik (2007) we get

$$\begin{aligned}
&\int_0^1 t^{\alpha-1}(1-t)^{\beta-1} B_t(\alpha+1,\beta)dt \\
&= \tfrac{1}{\alpha+1} \int_0^1 t^{2\alpha}(1-t)^{2\beta-1} \ {}_2F_1(\alpha+\beta+1,1;\alpha+2;t)dt \\
&= \tfrac{\alpha}{\alpha+\beta} \tfrac{B(2\alpha,2\beta)}{\alpha+1} \ {}_3F_2(\alpha+\beta+1,1,2\alpha+1;\alpha+2,2\alpha+2\beta+1;1)
\end{aligned}$$

Equation (B.15) in Milgram (2010) gives an explicit formula for our hypergeometric function in terms of gamma functions,

$$
\begin{aligned}
&{}_3F_2(\alpha+\beta+1, 1, 2\alpha+1; \alpha+2, 2\alpha+2\beta+1; 1) \\
&= \frac{\Gamma(\alpha+2)\Gamma(\beta)\Gamma(\alpha+\beta+0.5)2^{4\alpha+2\beta+1}}{\alpha\Gamma(0.5)\Gamma(2\alpha+2)\Gamma(2\alpha+2\beta+1)} \\
&\quad \times \left(\frac{(2\alpha+1)(2\alpha+2\beta)\Gamma(\alpha+\beta+0.5)\Gamma(\alpha+1)}{4\Gamma(\beta+0.5)} - \frac{\Gamma(\alpha+1.5)\Gamma(\alpha+\beta+1)}{\Gamma(0.5)\Gamma(\beta)} \right),
\end{aligned}
$$

which can be simplified using the duplication formula into

$$
\begin{aligned}
&\frac{2B(2\alpha, 2\beta)}{(\alpha+1)B(\alpha,\beta)^2}\ {}_3F_2(\alpha+\beta+1, 1, 2\alpha+1; \alpha+2, 2\alpha+2\beta+1; 1) \\
&= 1 - \frac{2B(2\alpha, 2\beta)}{\alpha B(\alpha,\beta)^2}
\end{aligned}
$$

for the end result of

$$
\mathbb{E}_{F_{\alpha,\beta}}|Y - Y'| = \tfrac{4B(2\alpha,2\beta)}{(\alpha+\beta)B(\alpha,\beta)^2}.
$$

A.2 Student's t-distribution

Proof To prove identity (4), we have

$$
\begin{aligned}
\mathbb{E}_{F_{\nu,\mu,\sigma}}|Y - y| &= \int_y^\infty (s-y)\tfrac{1}{\sigma} f_\nu\left(\tfrac{s-\mu}{\sigma}\right) ds - \int_{-\infty}^y (s-y)\tfrac{1}{\sigma} f_\nu\left(\tfrac{s-\mu}{\sigma}\right) ds \\
&= \int_{\frac{y-\mu}{\sigma}}^\infty (\sigma t - y + \mu) f_\nu(t) dt - \int_{-\infty}^{\frac{y-\mu}{\sigma}} (\sigma t - y + \mu) f_\nu(t) dt \\
&= (y-\mu)(2F_\nu(\tfrac{y-\mu}{\sigma}) - 1) + \sigma \left(\int_{\frac{y-\mu}{\sigma}}^\infty t f_\nu(t) dt - \int_{-\infty}^{\frac{y-\mu}{\sigma}} t f_\nu(t) dt \right).
\end{aligned}
$$

It is easily shown that

$$
G_\nu(x) = \int_{-\infty}^x t f_\nu(t) dt = -\tfrac{\nu}{\nu-1}\left(1 + \tfrac{x^2}{\nu}\right) f_\nu(x)
$$

which gives us

$$
\int_{\frac{y-\mu}{\sigma}}^\infty t f_\nu(t) dt = \frac{\nu + (\frac{y-\mu}{\sigma})^2}{\nu - 1} f_\nu(\tfrac{y-\mu}{\sigma}) = -\int_{-\infty}^{\frac{y-\mu}{\sigma}} t f_\nu(t) dt
$$

For identity (5), we have

$$
\begin{aligned}
\mathbb{E}_{F_{\nu,\mu,\sigma}}|Y-Y'| &= 2\int_0^\infty s' \ (f_{\nu,\sigma} * f_{\nu,\sigma})(s')ds' \\
&= 2\int_0^\infty \int_{-\infty}^\infty \frac{s'}{\sigma^2} f_\nu(\tfrac{t'}{\sigma}) f_\nu(\tfrac{s'-t'}{\sigma})dt'ds' \\
&= 2\sigma \int_{-\infty}^\infty f_\nu(t)\int_{-t}^\infty (t+s)f_\nu(s)dsdt \\
&= 2\sigma\left[\int_{-\infty}^\infty tf_\nu(t)(1-F_\nu(-t))dt - \int_{-\infty}^\infty f_\nu(t)G_\nu(-t)dt\right] \\
&= 2\sigma\left[\Big[G_\nu(t)(1-F_\nu(-t))\Big]_{-\infty}^\infty - \int_{-\infty}^\infty G_\nu(t)f_\nu(-t)dt - \int_{-\infty}^\infty f_\nu(t)G_\nu(-t)dt\right] \\
&= -4\sigma\int_{-\infty}^\infty f_\nu(t)G_\nu(t)dt \\
&= \frac{4\sigma}{(\nu-1)B(\frac{1}{2},\frac{\nu}{2})^2}\int_{-\infty}^\infty \left(1+\frac{t^2}{\nu}\right)^{-\nu} dt.
\end{aligned}
$$

The remaining integral can be calculated using the substitution $u = \left(1+\frac{t^2}{\nu}\right)^{-1}$,

$$
\begin{aligned}
\int_{-\infty}^\infty \left(1+\frac{t^2}{\nu}\right)^{-\nu} dt &= 2\int_0^\infty \left(1+\frac{t^2}{\nu}\right)^{-\nu} dt \\
&= \sqrt{\nu}\int_0^1 (1-u)^{-\frac{1}{2}}u^{\nu-\frac{3}{2}} \ du \\
&= \sqrt{\nu}B(\tfrac{1}{2},\nu-\tfrac{1}{2}).
\end{aligned}
$$