# Forecast Similarity Using Cramer Distance Approximation

Johannes Bracher, Evan Ray, Nick Reich, Nutcha Wattanachit

07/15/2021

## Cramer Distance

Consider two predictive distributions $F$ and $G$. Their *Cramer distance* or *integrated quadratic distance* is defined as

$$\text{CD}(F, G) = \int_{-\infty}^{\infty} (F(x) - G(x))^2 dx$$

where $F(x)$ and $G(x)$ denote the cumulative distribution functions. The Cramer distance is the divergence associated with the continuous ranked probability score (Thorarinsdottir 2013, Gneiting and Raftery 2007), which is defined by

$$\text{CRPS}(F, y) = \int_{-\infty}^{\infty} (F(x) - \mathbf{1}(x \geq y))^2 dx =) = 2 \int_0^1 ((\mathbf{1}(y \leq q_k^F) - \tau_k)(q_k^F - y) d\tau_k \tag{1}$$

$$\text{CD}(F, G) = \mathbb{E}_{F,G}|x - y| - 0.5 \left[\mathbb{E}_F|x - x'| + \mathbb{E}_G|y - y'|\right], \tag{2}$$

where $x, x'$ are independent random variables following $F$ and $y, y'$ are independent random variables following $G$. This formulation illustrates that the Cramer distance depends on the shift between $F$ and $G$ (first term) and the variability of both $F$ and $G$ (of which the two last expectations in above equation are a measure).

### Cramer Distance Approximation for Equally-Spaced Intervals

Now assume that for each of the distributions $F$ and $G$ we only know $K$ quantiles at equally spaced levels $1/(K + 1), 2/(K + 1), \ldots, K/(K + 1)$. Denote these quantiles by $q_1^F, \ldots, q_K^F$ and $q_1^G, \ldots, q_K^G$, respectively. It is well known that the CRPS can be approximated by an average of linear quantile scores (Laio and Tamea 2007, Gneiting and Raftery 2007):

$$\text{CRPS}(F, y) \approx \frac{1}{K} \times \sum_{k=1}^{K} 2\{\mathbf{1}(y \leq q_k^F) - \tau_k\} \times (q_k^F - y). \tag{3}$$

This approximation is equivalent to the weighted interval score (WIS) which is in use for evaluation of quantile forecasts at the Forecast Hub, see Section 2.2 of Bracher et al (2021). This approximation can be generalized to the Cramer distance as

$$\text{CD}(F, G) \approx \frac{1}{K(K + 1)} \times \sum_{i=1}^{2K-1} b_i(b_i + 1)(q_{i+1} - q_i), \tag{4}$$

where we use the following notation:

- $\mathbf{q}$ is a vector of length $2K$. It is obtained by pooling the $q_k^F, q_k^G, k = 1, \dots, K$ and ordering them in increasing order (ties can be ordered in an arbitrary manner).
- $\mathbf{a}$ is a vector of length $2K$ containing the value 1 wherever $\mathbf{q}$ contains a quantile of $F$ and $-1$ wherever it contains a value of $G$.
- $\mathbf{b}$ is a vector of length $2K$ containing the absolute cumulative sums of $\mathbf{a}$, i.e. $b_i = \left|\sum_{j=1}^{i} a_j\right|$.

For small $K$ it seems that the slightly different approximation

$$\mathrm{CD}(F, G) \approx \frac{1}{(K+1)^2} \times \sum_{i=1}^{2K-1} b_i^2 (q_{i+1} - q_i), \tag{5}$$

actually works better. This just corresponds to the integrated squared difference between two step functions $F^*$ and $G^*$ with $F^*(x) = 0$ for $x < q_1^F$, $F^*(x) = k/(K+1)$ for $q_k^F \leq x < q_k^F$, $F^*(x) = K/(K+1)$ for $x \geq x_K^F$ and $G^*$ defined accordingly. We illustrate this in the figure below, with light blue areas representing the CD and approximated CD.

## Cramer Distance Approximation for Unequally-Spaced Intervals

Suppose we have quantiles $q_1^F, \dots, q_K^F$ and $q_1^G, \dots, q_K^G$ at $K$ probability levels $\tau_1, \dots, \tau_K$ from two distributions $F$ and $G$. Define the combined vector of quantiles $q_1, \dots, q_{2K}$ by combining the vectors $q_1^F, \dots, q_K^F$ and $q_1^G, \dots, q_K^G$ and sorting them in an ascending order. The CRPS can be approximated as follows

$$\mathrm{CRPS}(F, y) \approx 2 \sum_{k=1}^{K} \{\mathbf{1}(y \leq q_k^F) - \tau_k\} \times (q_k^F - y) \times (\tau_k - \tau_{k-1}). \tag{6}$$

with $\tau_0 = 0$. We can see that with equally spaced interval of $1/K$ increment, this equation generalizes to (**??**).

Essentially, we can approximate the Cramer distance by eliminating the tails of the integral to the left of $q_1$ and the right of $q_{2K}$, and approximating the center via a Riemann sum:

$$\begin{aligned}
\mathrm{CD}(F, G) &= \int_{-\infty}^{\infty} F(x) - G(x)^2 dx && (7)\\
&\approx \int_{q_1}^{q_{2K}} F(x) - G(x)^2 dx && (8)\\
&= \sum_{j=1}^{2K-1} \int_{q_j}^{q_{j+1}} F(x) - G(x)^2 dx && (9)
\end{aligned}$$

There are a variety of options that can be used for each term in this sum, for instance:

### Left-sided Riemann sum approximation

$$\begin{aligned}
\mathrm{CD}(F, G) &\approx \sum_{j=1}^{2K-1} \int_{q_j}^{q_{j+1}} F(x) - G(x)^2 && (10)\\
&\approx \sum_{j=1}^{2K-1} \{\hat{F}(q_j) - \hat{G}(q_j)\}^2 (q_{j+1} - q_j) && (11)\\
& && (12)
\end{aligned}$$

Since $q_j \in \{q_1, ..., q_{2K}\}$ belongs to either $q_1^F, ..., q_K^F$ or $q_1^G, ..., q_K^G$, we can rewrite the above approximation using $\tau_1, ..., \tau_K$ as follows

$$\mathrm{CD}(F, G) \approx \sum_{j=1}^{2K-1} \{\hat{F}(q_j) - \hat{G}(q_j)\}^2 (q_{j+1} - q_j) \tag{13}$$

$$= \sum_{j=1}^{2K-1} \{\tau_j^F - \tau_j^G\}^2 (q_{j+1} - q_j) \tag{14}$$

where $\tau_j^F \in \tau_F$ and $\tau_j^G \in \tau_G$. $\tau_F$ and $\tau_G$ are vectors of length $2K-1$ with elements

$$\tau_j^F = \begin{cases} I(q_1 = q_1^F) \times \tau_{q_1}^F & \text{for } j = 1 \\ I(q_j \in \{q_1^F, ..., q_K^F\}) \times \tau_{q_j}^F + I(q_j \in \{q_1^G, ..., q_K^G\}) \times \tau_{j-1}^F & \text{for } j > 1 \end{cases}$$

where $\tau_{q_j}^F$ is the probability level corresponding to $q_j$ given $q_j$ in the pooled quantiles comes from $F$, and $\tau_{j-1}^F$ is the $(j-1)^{th}$ probability level in $\tau_F$.

$$\tau_j^G = \begin{cases} I(q_1 = q_1^G) \times \tau_{q_1}^G & \text{for } j = 1 \\ I(q_j \in \{q_1^G, ..., q_K^G\}) \times \tau_{q_j}^G + I(q_j \in \{q_1^F, ..., q_K^F\}) \times \tau_{j-1}^G & \text{for } j > 1 \end{cases}$$

where $\tau_{q_j}^G$ is the probability level corresponding to $q_j$ given $q_j$ in the pooled quantiles comes from $G$, and $\tau_{j-1}^G$ is the $(j-1)^{th}$ probability level in $\tau_G$.

**Trapezoidal rule**

$$\mathrm{CD}(F, G) \approx \sum_{j=1}^{2K-1} \int_{q_j}^{q_{j+1}} F(x) - G(x)^2 \tag{15}$$

$$\approx \sum_{j=1}^{2K-1} \frac{\{\hat{F}(q_j) - \hat{G}(q_j)\}^2 + \{\hat{F}(q_{j+1}) - \hat{G}(q_{j+1})\}^2}{2} (q_{j+1} - q_j) \tag{16}$$

$$\tag{17}$$

Similarly, we can rewrite the above approximation using $\tau_1, ..., \tau_K$ as defined in the left-sided Riemann sum approximation as follows

$$\mathrm{CD}(F, G) \approx \sum_{j=1}^{2K-1} \frac{\{\hat{F}(q_j) - \hat{G}(q_j)\}^2 + \{\hat{F}(q_{j+1}) - \hat{G}(q_{j+1})\}^2}{2} (q_{j+1} - q_j) \tag{18}$$

$$= \sum_{j=1}^{2K-1} \frac{\{\tau_j^F - \tau_j^G\}^2 + \{\tau_{j+1}^F - \tau_{j+1}^G\}^2}{2} (q_{j+1} - q_j). \tag{19}$$

## Cramer Distance Approximation for Unequally-Spaced Intervals and Different Probability Levels

We (probably) can further modify the formula of the Cramer distance approximation for unequally-spaced intervals to accommodate different probability levels from $F$ and $G$. Suppose we have quantiles $q_1^F, ..., q_N^F$ at

$K$ probability levels $\tau_1^F, ..., \tau_N^F$ from the distribution $F$, and $q_1^G, ..., q_M^G$ at $M$ probability levels $\tau_1^G, ..., \tau_M^G$ from the distribution $G$. Define the combined vector of quantiles $q_1, ..., q_{N+M}$ by combining the vectors $q_1^F, ..., q_N^F$ and $q_1^G, ..., q_M^G$ and again sorting them in an ascending order. Using the same definitions as previously defined, we can approximate the Cramer distance via a Riemann sum as follows:

**Left-sided Riemann sum approximation**

$$\mathrm{CD}(F,G) \approx \sum_{j=1}^{N+M-1} \int_{q_j}^{q_{j+1}} F(x) - G(x)^2 \tag{20}$$

$$\approx \sum_{j=1}^{N+M-1} \{\hat{F}(q_j) - \hat{G}(q_j)\}^2 (q_{j+1} - q_j), \tag{21}$$

$$\tag{22}$$

which we can rewrite using $\tau_1^F, ..., \tau_N^F$ and $\tau_1^G, ..., \tau_M^G$ as follows

$$\mathrm{CD}(F,G) \approx \sum_{j=1}^{N+M-1} \{\hat{F}(q_j) - \hat{G}(q_j)\}^2 (q_{j+1} - q_j) \tag{23}$$

$$= \sum_{j=1}^{N+M-1} \{\tau_j^F - \tau_j^G\}^2 (q_{j+1} - q_j) \tag{24}$$

where $\tau_j^F \in \tau_F$ and $\tau_j^G \in \tau_G$. $\tau_F$ and $\tau_G$ are vectors of length $N + M - 1$ with elements

$$\tau_j^F = \begin{cases} \tau_{q_j}^F & \text{if } q_j \in \{q_1^F, ..., q_N^F\} \\ \tau_{q_{j-1}}^F & \text{if } q_j \notin \{q_1^F, ..., q_N^F\} \end{cases}$$

where $\tau_{q_j}^F$ is the probability level corresponding to $q_j$ given $q_j$ in the pooled quantiles comes from $F$.

$$\tau_j^G = \begin{cases} \tau_{q_j}^G & \text{if } q_j \in \{q_1^G, ..., q_M^G\} \\ \tau_{q_{j-1}}^G & \text{if } q_j \notin \{q_1^G, ..., q_M^G\} \end{cases}$$

where $\tau_{q_j}^G$ is the probability level corresponding to $q_j$ given $q_j$ in the pooled quantiles comes from $G$.

**Trapezoidal rule**

$$\mathrm{CD}(F,G) \approx \sum_{j=1}^{2K-1} \int_{q_j}^{q_{j+1}} F(x) - G(x)^2 \tag{25}$$

$$\approx \sum_{j=1}^{N+M-1} \frac{\{\hat{F}(q_j) - \hat{G}(q_j)\}^2 + \{\hat{F}(q_{j+1}) - \hat{G}(q_{j+1})\}^2}{2} (q_{j+1} - q_j), \tag{26}$$

$$\tag{27}$$

which we can rewrite as follows

$$\mathrm{CD}(F,G) \approx \sum_{j=1}^{N+M-1} \frac{\{\hat{F}(q_j) - \hat{G}(q_j)\}^2 + \{\hat{F}(q_{j+1}) - \hat{G}(q_{j+1})\}^2}{2} (q_{j+1} - q_j) \tag{28}$$

$$= \sum_{j=1}^{N+M-1} \frac{\{\tau_j^F - \tau_j^G\}^2 + \{\tau_{j+1}^F - \tau_{j+1}^G\}^2}{2} (q_{j+1} - q_j). \tag{29}$$

## Decomposition of Approximated Cramer Distance

The Cramer distance is commonly used to measure the similarity of forecast distributions (see Richardson et al 2020 for a recent application). Now assume that for each of the distributions $F$ and $G$ we only know $K$ quantiles at equally spaced levels $1/(K+1), 2/(K+1), \dots, K/(K+1)$. Denote these quantiles by $q_1^F, \dots, q_K^F$ and $q_1^G, \dots, q_K^G$, respectively. This CRPS approximation given by (**??**) is equivalent to the weighted interval score (WIS) which is in use for evaluation of quantile forecasts at the Forecast Hub, see Section 2.2 of Bracher et al (2021). This approximation can be generalized to the Cramer distance as

$$\mathrm{CD}(F,G) \approx \frac{1}{K(K+1)} \sum_{i=1}^{K} \sum_{j=1}^{K} \mathbf{1}\{(i-j) \times (q_i^F - q_j^G) \leq 0\} \times \left| q_i^F - q_j^G \right|, \tag{30}$$

This can be seen as a sum of penalties for *incompatibility* of predictive quantiles. Whenever the predictive quantiles $q_i^F$ and $q_j^G$ are incompatible in the sense that they imply $F$ and $G$ are different distributions (e.g. because $q_F^i > q_G^j$ despite $i < j$ or $q_F^i \neq q_G^j$ despite $i = j$), a penalty $\left| q_i^F - q_j^G \right|$ is added to the sum. This corresponds to the shift which would be necessary to make $q_F^i$ and $q_G^j$ compatible.

## A divergence measure for central prediction intervals with potentially different nominal coverages

Consider two central prediction intervals $[l^F, u^F]$ and $[l^G, u^G]$ with nominal levels $\alpha^F$ and $\alpha^G$, respectively (meaning that $l^F$ is the $(1-\alpha^F)/2$ quantile of $F$ etc). We can define an *interval divergence* measure by comparing the two pairs of predictive quantiles and summing up the respective incompatibility penalties as in (30). Adapting notation to the interval formulation and structuring the sum slightly differently, this can be written as:

$$\begin{aligned}
\mathrm{ID}([l^F, u^F], [l^G, u^G], \alpha^F, \alpha^G) = \; & \mathbf{1}(\alpha^F \leq \alpha^G) \times \{\max(l^G - l^F, 0) + \max(u^F - u^G, 0)\} \; + \\
& \mathbf{1}(\alpha^F \geq \alpha^G) \times \{\max(l^F - l^G, 0) + \max(u^G - u^F, 0)\} \; + \\
& \max(l^F - u^G, 0) \; + \\
& \max(l^G - u^F, 0)
\end{aligned}$$

The first row adds penalties for the case where $[l^F, u^F]$ should be nested in $[l^G, u^G]$, but at least one of its ends is more extreme than the respective end of $[l^G, u^G]$. The second row covers the converse case. The last two rows add penalties if the lower end of one interval exceeds the upper end of the other, i.e. the intervals do not overlap.

This can be seen as a (scaled version of a) generalization of the interval score, but writing out the exact relationship is a bit tedious.

We now define four auxiliary terms with an intuitive interpretation which add up to the interval divergence:

- The term
$$D_F = \mathbf{1}(\alpha^F \leq \alpha^G) \times \max\{(u^F - l^F) - (u^G - l^G), 0\}$$
is the sum of penalties resulting from $F$ being more dispersed than $G$. It is positive whenever the interval $[l^F, u^F]$ is longer than $[l^G, u^G]$, even though it should be nested in the latter. $D_F$ then tells us by how much we would need to shorten $[l^F, u^F]$ so it could fit into $[l^G, u^G]$.
- The term
$$D_G = \mathbf{1}(\alpha^G \leq \alpha^G) \times \max\{(u^G - l^G) - (u^F - l^F), 0\}$$
measures the converse, i.e. overdispersion of $G$ relative to $F$.
- The term
$$S^F = \max\{\mathbf{1}(\alpha^G \leq \alpha^F)\times\max(l^F - l^G, 0) + \mathbf{1}(\alpha^F \leq \alpha^G)\times\max(u^F - u^G, 0) + \max(l^F - u^G, 0) - D_F - D_G, 0\}$$
sums over penalties for values in $\{l^F, u^F\}$ exceeding those from $\{l^G, u^G\}$ where they should not (only counting penalties not already covered in $D_F$ or $D_G$). It thus represents an *upward shift* of $F$ relative to $G$.
- The term
$$S^G = \max\{\mathbf{1}(\alpha^F \leq \alpha^G)\times\max(l^G - l^F, 0) + \mathbf{1}(\alpha^G \leq \alpha^F)\times\max(u^G - u^F, 0) + \max(l^G - u^F, 0) - D_G - D_F, 0\}$$
accordingly represents an *upward shift* of $G$ relative to $F$.

It can be shown that
$$\mathrm{ID}([l^F, u^F], [l^G, u^G], \alpha^F, \alpha^G) = D_F + D_G + S^F + S^G$$
Intuitively the interval divergence measures by how much we need to move the quantiles of the interval with lower nominal coverage so it fits into the one with larger nominal coverage.

## Approximating the Cramer distance using interval divergences

Assuming $K$ is even, the $K$ equally spaced predictive quantiles of each distribution can seen as $L = K/2$ central prediction intervals with coverage levels $\alpha_i = 2i/(L+1), i = 1, \dots, L$. Similarly to the definition of the WIS, the approximation (30) can also be expressed in terms of these intervals as

$$\mathrm{CD}(F, G) \approx \frac{1}{2L(2L+1)} \sum_{k=1}^{L} \sum_{m=1}^{L} \mathrm{ID}([l_k^F, u_k^F], [l_m^G, u_m^G], \alpha_k^F, \alpha_m^G).$$

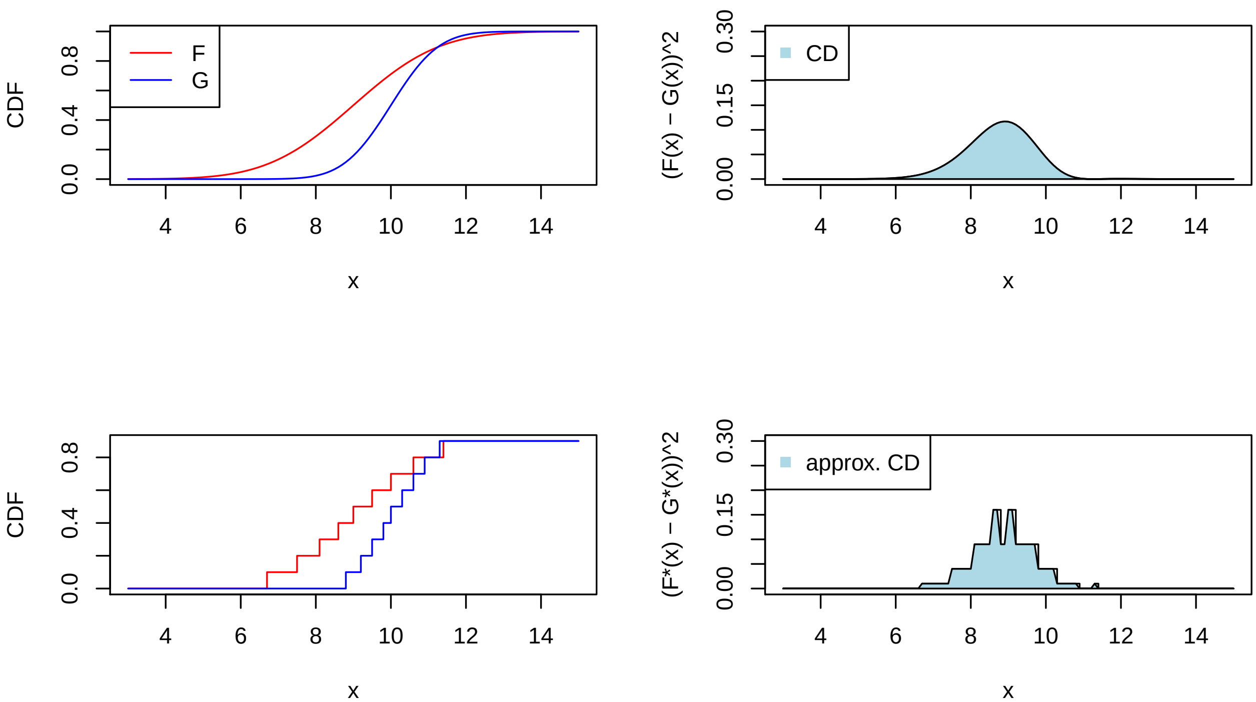
This implies a decomposition of the Cramer distance into the four interpretable components defined for the interval divergence in the previous section. If $G$ is a one-point distribution, the CD reduces to the WIS and the proposed decomposition reduces to the well-known decomposition of the WIS into dispersion, overprediction and underprediction components.

Note that in practice we usually have an uneven rather than even number $K$ of predictive quantiles. In this case the median needs to be treated separately (comparisons of the "0% prediction interval'' need to be weighted down with a factor of 2; this is the same little quirk as the one identified by Ryan and Evan for the WIS a few months ago). The decomposition has the following properties:

- Additive shifts of the two distributions only affect the shift components, not the dispersion components.
- Consequently, if $G$ and $G$ are identical up to an additive shift, both dispersion components will be 0.
- If $F$ and $G$ are both symmetric and have the same median, the both shift components will be 0.
- I think that in general it is possible that both shift components or both dispersion components are greater than 0, which leads to a somewhat strange interpretation. But this should only concern constructed examples.

# Examples

## Equally-spaced intervals




In this example, six different approximations are applied to the distributions $F\ N(9, 1.8)$ and $G\ N(10, 1)$ in the figures above.

- Using direct numerical integration based on a fine grid of values for $x$:

```
## [1] 0.2532376
```

- Using sampling and the alternative expression (2) of the CD from above:

```
## [1] 0.2457156
```

- Using the first quantile-based approximation (5) and various values of $K$:

```
## [1] 0.3550788 0.3078906 0.2764153 0.2652018 0.2593619 0.2557450 0.2545077
## [8] 0.2538792
```

- Using the second quantile-based approximation (5) and various values of $K$:

```
## [1] 0.2926809 0.2723571 0.2608768 0.2572045 0.2552998 0.2541028 0.2536835
## [8] 0.2534662
```

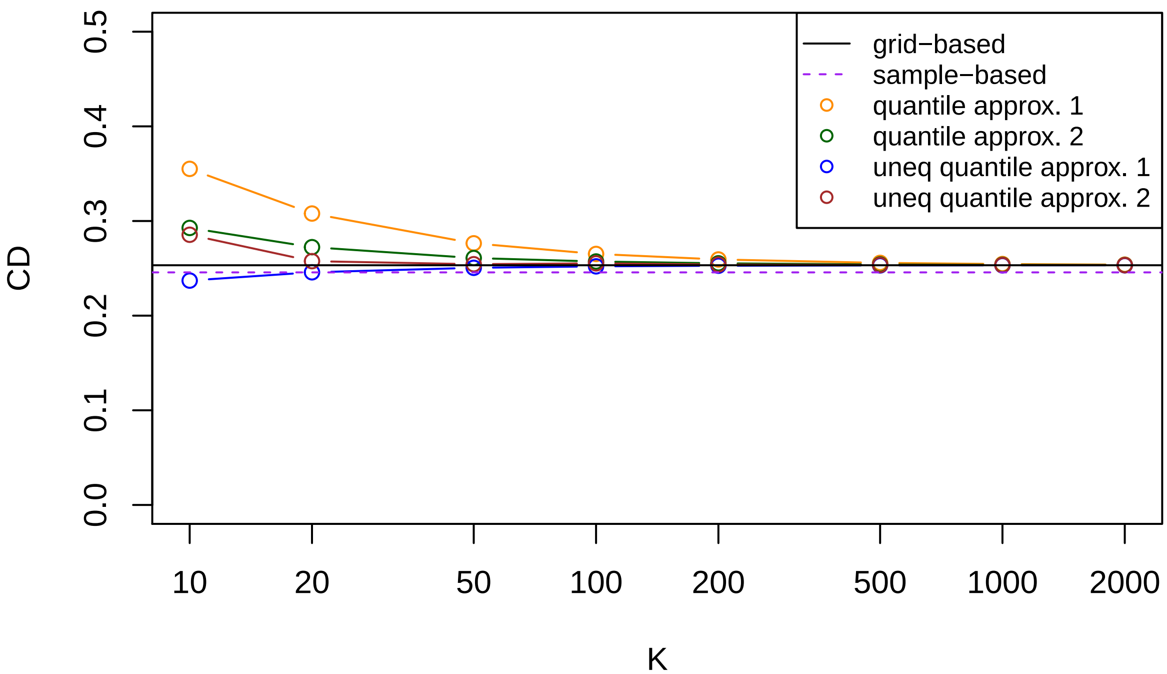
- Using the left-sided Riemann sum approximation and various values of $K$:

```
## [1] 0.2370715 0.2458022 0.2505461 0.2520862 0.2527531 0.2530874 0.2531764
## [8] 0.2532128
```

- Using the trapezoidal Riemann sum approximation and various values of $K$:

```
## [1] 0.2854597 0.2575762 0.2543386 0.2552775 0.2540318 0.2535609 0.2534094
## [8] 0.2533309
```

The below plot shows the results from the different computations.




In the case that $G$ is a point mass at $y = 10$, approximation (4) indeed coincides with (3).

```
## [1] "Quantile approx. 1: 0.688567227886639"
## [1] "Quantile approx. 2: 0.608983067073759"
## [1] "Uneq quantile approx. 1: 1.03814992169128"
## [1] "Uneq quantile approx. 2: 1.24791020451193"
## [1] "Quantile score WIS: 0.688567227886639"
```

The approximation (4) is closer to the grid-based direct evaluation of the integral. Since the unequally-spaced approximations were not formulated from (equally-spaced) WIS, it may be expected.

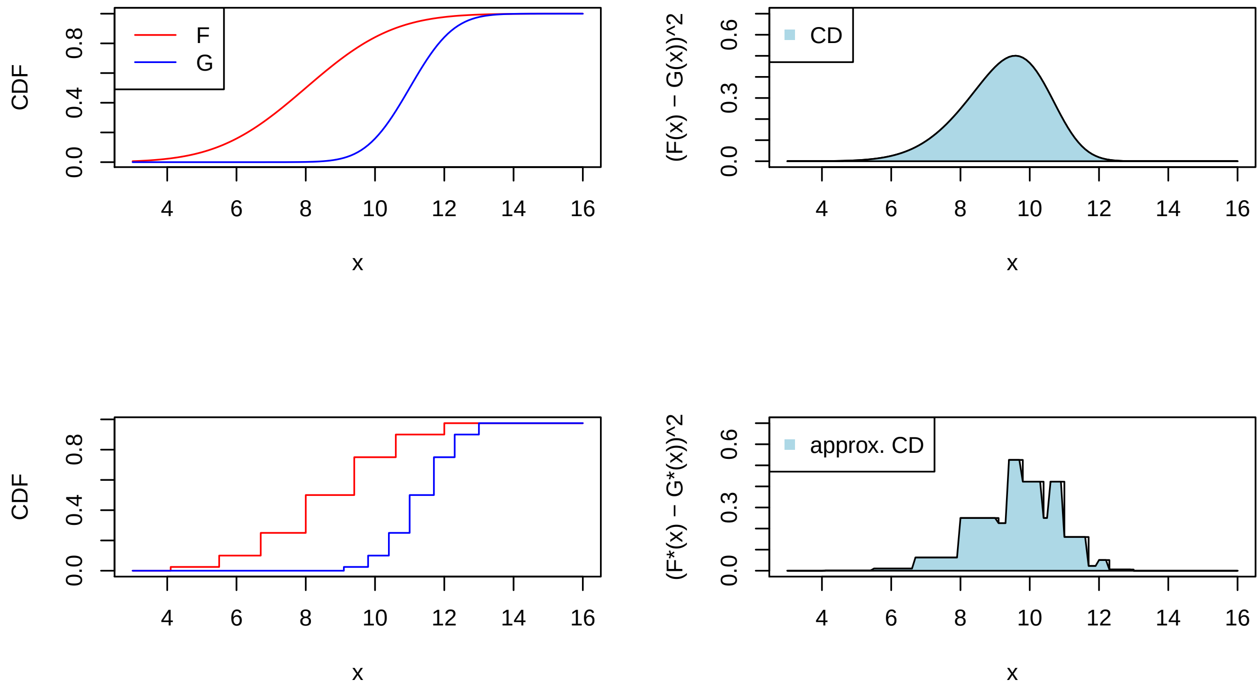
```
## [1] "Grid-based approx.: 0.61599852942592"
```

## Unequally-spaceed intervals

We apply the same six approximations as in the previous example to the two distributions $F \sim N(8, 2)$ and $G \sim N(11, 1)$ whose quantiles correspond to unequally-spaced probability levels.

### 7 quantiles with unequally-spaced intervals

The probability levels corresponding to the given set of quantiles in this example is $0.025, 0.1, 0.25, 0.5, 0.75, 0.9, 0.975$, which is the same probability levels provided by the COVID-hub case forecasts.




- Using direct numerical integration based on a fine grid of values for $x$.

```
## [1] 1.493653
```

- Using sampling and the alternative expression (2) of the CD from above:

```
## [1] 1.483948
```

- Using the first quantile-based approximation:

```
## [1] 1.919252
```

- Using the second quantile-based approximation:

```
## [1] 1.764859
```

- Using the left-sided Riemann sum-based approximation:

```
## [1] 1.35122
```

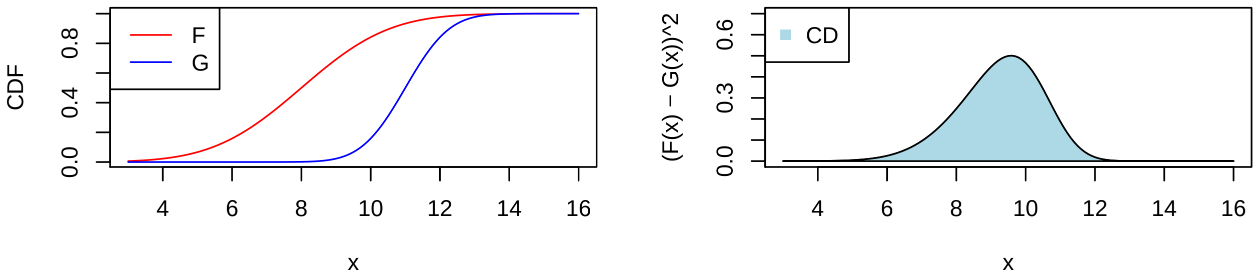
- Using the trapezoidal Riemann sum-based approximation:

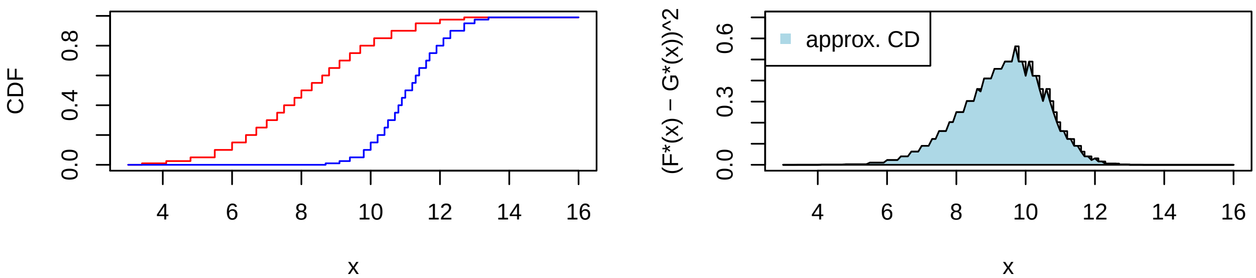
```
## [1] 1.468801
```

Out of all four quantile-based approximation, the trapezoidal Riemann sum-based approximation is closest to the grid-based integral evaluation.

**23 quantiles with 2 unequally-spaced probability levels at the tails**

Using the same $F$ and $G$, the probability levels corresponding to the given set of quantiles in this example is the same probability levels provided by the COVID-hub death forecasts. They are almost equally-spaced, except at the tails.




- Using the first quantile-based approximation:

```
## [1] 1.640408
```

- Using the second quantile-based approximation:

```
## [1] 1.581296
```

- Using the left-sided Riemann sum-based approximation:

```
## [1] 1.452266
```

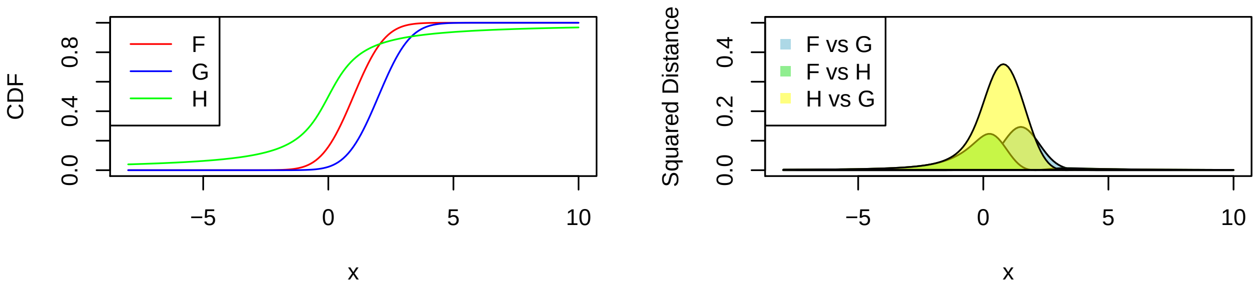
- Using the trapezoidal Riemann sum-based approximation:

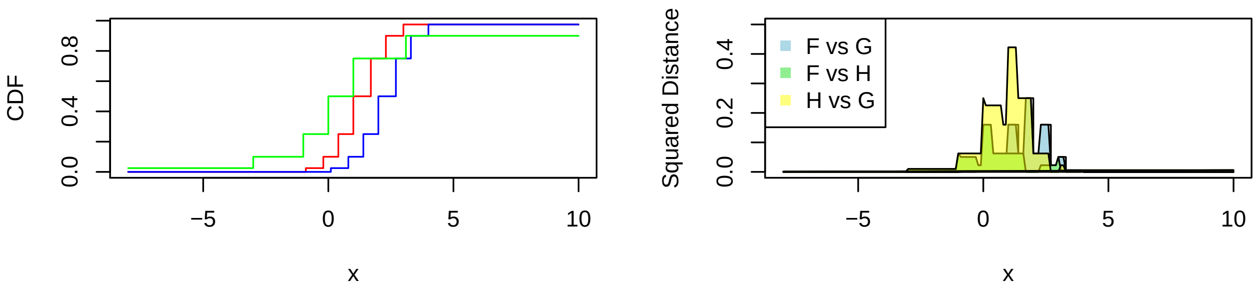
```
## [1] 1.470718
```

Again, the trapezoidal Riemann sum-based approximation is closest to the grid-based integral evaluation of 1.493653.

### Examples of Disagreement Between Equally- and Unequally-spaced Interval Methods

**Heavy tails** Suppose we have three cumulative distributions, $F \sim N(1,1)$, $G \sim N(2,1)$ and $H \sim T_1$, represented by 7 unequally-spaced quantiles. The probability levels corresponding to the given set of quantiles in this example is $0.025, 0.1, 0.25, 0.5, 0.75, 0.9, 0.975$.




- Using direct numerical integration based on a fine grid of values for $x$.

```
## [1] "CD of F vs G: 0.270903289652979"
## [1] "CD of F vs H: 0.303008857878541"
## [1] "CD of H vs G: 0.830227986212452"
```

- Using the first quantile-based approximation:

```
## [1] "Approx. CD of F vs G: 0.430834455349389"
## [1] "Approx. CD of F vs H: 0.412836097817344"
## [1] " Approx. CD of H vs G: 1.0891147790307"
```

- Using the second quantile-based approximation:

```
## [1] "Approx. CD of F vs G: 0.349525091827873"
## [1] "Approx. CD of F vs H: 0.318185573750552"
## [1] " Approx. CD of H vs G: 0.958988318892232"
```

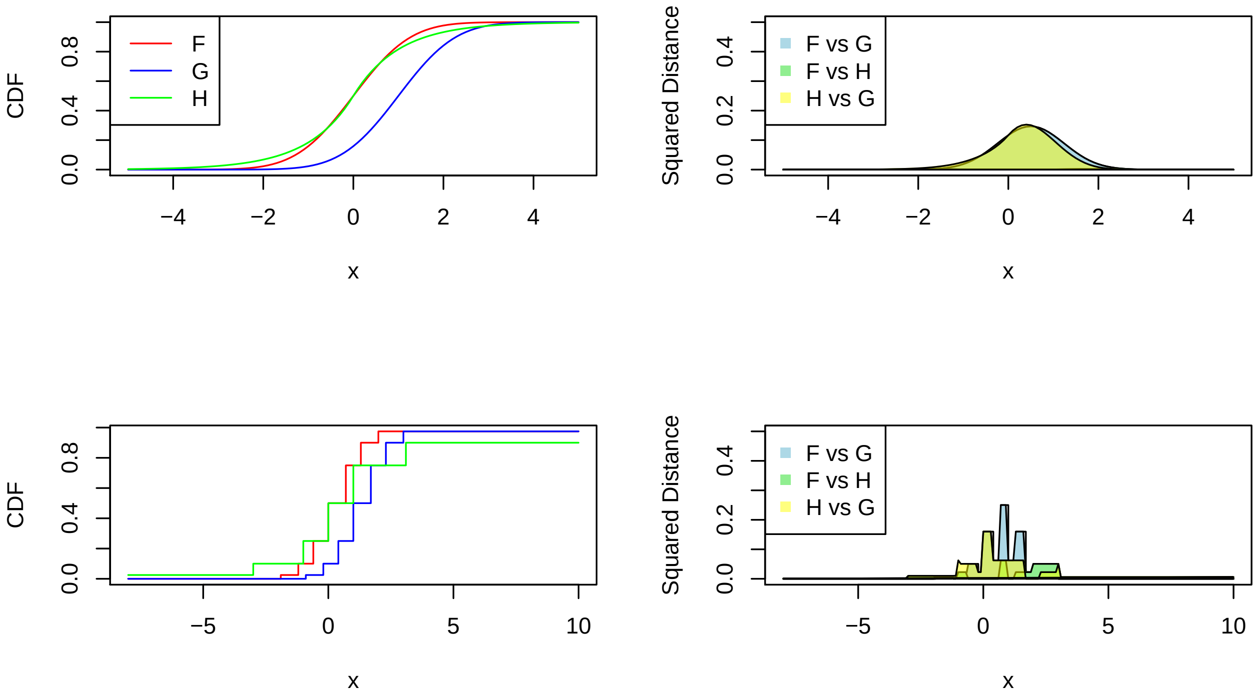
- Using the left-sided Riemann sum-based approximation:

```
## [1] "Approx. CD of F vs G: 0.267605148430715"
## [1] "Approx. CD of F vs H: 0.24361082990276"
## [1] " Approx. CD of H vs G: 0.734225431651865"
```

- Using the trapezoidal Riemann sum-based approximation:

```
## [1] "Approx. CD of F vs G: 0.302511061162121"
## [1] "Approx. CD of F vs H: 0.266926890705267"
## [1] " Approx. CD of H vs G: 0.752143988834321"
```

**Long tails** Suppose we have three cumulative distributions, $F \sim N(0,1)$, $G \sim N(1,1)$ and $H \sim \text{Laplace}(0,1)$, represented by 7 unequally-spaced quantiles. The probability levels corresponding to the given set of quantiles in this example is $0.025, 0.1, 0.25, 0.5, 0.75, 0.9, 0.975$.




- Using direct numerical integration based on a fine grid of values for $x$.

```
## [1] "CD of F vs G: 0.270903289581517"
## [1] "CD of F vs H: 0.0068412250422997"
## [1] "CD of H vs G: 0.257655267503305"
```

- Using the first quantile-based approximation:

```
## [1] "Approx. CD of F vs G: 0.430834455349389"
## [1] "Approx. CD of F vs H: 0.0203971216157386"
## [1] " Approx. CD of H vs G: 0.416591384312851"
```

- Using the second quantile-based approximation:

```
## [1] "Approx. CD of F vs G: 0.349525091827873"
## [1] "Approx. CD of F vs H: 0.0116554980661363"
## [1] " Approx. CD of H vs G: 0.333247296357544"
```

- Using the left-sided Riemann sum-based approximation:

```
## [1] "Approx. CD of F vs G: 0.267605148430715"
## [1] "Approx. CD of F vs H: 0.00892374070688564"
## [1] " Approx. CD of H vs G: 0.255142461273745"
```

- Using the trapezoidal Riemann sum-based approximation:

```
## [1] "Approx. CD of F vs G: 0.302511061162121"
## [1] "Approx. CD of F vs H: 0.0194133332014688"
```

```
## [1] " Approx. CD of H vs G: 0.259617817413175"
```