JQuantile

The class JQuantile resides in the JTOOLS namespace and constitutes a kind of zero-dimensional histogram. It contains a running mean and standard deviation and can optionally be used to determine quantiles.

The following code extract:

```
const double x     = 0.0;
const double sigma = 1.0;

JQuantile quantile;

for (int i = 0; i != 1000; ++i) {
  quantile.put(gRandom->Gaus(x, sigma));
}

NOTICE("quantity  " << CENTER(10)  <<   "calculated"     << " | " << CENTER(10)  << "true" << endl);
NOTICE("mean      " << FIXED(10,3) << quantile.getMean()  << " | " << FIXED(10,3) <<  x     << endl);
NOTICE("RMS       " << FIXED(10,3) << quantile.getSTDev() << " | " << FIXED(10,3) <<  sigma << endl);
```

could produce:

```
quantity  calculated |    true
mean          -0.006 |      0.000
RMS            1.024 |      1.000
```

Recurrence relations exists for the running mean and standard deviation [1]. Here, the recurrence relations for the weighed mean and standard deviation are presented.

For a data sample $\{(w_1, x_1), \ldots, (w_n, x_n)\}$, the following definitions are adopted.

$$W_n \equiv \sum_{i=1}^{n} w_i \tag{1}$$

$$\mu_n \equiv \frac{1}{W_n} \sum_{i=1}^{n} w_i x_i \tag{2}$$

$$\sigma_n^2 \equiv \frac{1}{\langle W_n \rangle (n-1)} \sum_{i=1}^{n} w_i (x_i - \mu_n)^2 \tag{3}$$

where W_n is the total weight, μ_n the weighed mean value and σ_n^2 the square of the weighed standard deviation. In this, $\langle W_n \rangle$ corresponds to the average weight $\langle W_n \rangle \equiv \frac{W_n}{n}$.

We start with the derivation of the recurrence relation for the weighed mean value μ.

$$\mu_n = \frac{1}{W_n} \sum_{i=1}^{n} w_i x_i \tag{4}$$

$$= \frac{1}{W_n} \left[\sum_{i=1}^{n-1} w_i x_i + w_n x_n \right] \tag{5}$$

$$= \frac{1}{W_n}\left[\mu_{n-1}W_{n-1} + w_n x_n\right] \quad (6)$$

$$= \frac{1}{W_n}\left[\mu_{n-1}W_n - \mu_{n-1}w_n + w_n x_n\right] \quad (7)$$

$$= \mu_{n-1} + \frac{w_n}{W_n}\left[x_n - \mu_{n-1}\right] \quad (8)$$

For the weighed standard deviation, we define first:

$$s_n^2 \equiv \langle W_n \rangle (n-1)\sigma_n^2 \quad (9)$$

It then follows that:

$$s_n^2 = \sum_{i=1}^{n} w_i x_i^2 - W_n \mu_n^2 \quad (10)$$

$$= \sum_{i=1}^{n-1} w_i x_i^2 + w_n x_n^2 - W_n\left[\mu_{n-1}^2 - 2\mu_{n-1}\frac{w_n}{W_n}(x_n - \mu_{n-1}) + \left(\frac{w_n}{W_n}(x_n - \mu_{n-1})\right)^2\right] \quad (11)$$

$$= \sum_{i=1}^{n-1} w_i x_i^2 + w_n x_n^2 - W_{n-1}\mu_{n-1}^2 - w_n\mu_{n-1}^2 - 2\mu_{n-1}w_n(x_n - \mu_{n-1}) - \frac{w_n^2(x_n - \mu_{n-1})^2}{W_n} \quad (12)$$

$$= s_{n-1}^2 + w_n(x_n - \mu_{n-1})^2 + w_n(\mu_n - \mu_{n-1})(x_n - \mu_{n-1}) \quad (13)$$

$$= s_{n-1}^2 + w_n(x_n - \mu_{n-1})(x_n - \mu_n) \quad (14)$$

Hence:

$$\sigma_n^2 = \frac{1}{\langle W_n \rangle (n-1)} s_n^2 \quad (15)$$

$$= \frac{1}{\langle W_n \rangle (n-1)}\left[s_{n-1}^2 + w_n(x_n - \mu_{n-1})(x_n - \mu_n)\right] \quad (16)$$

References

[1] Knuth TAOCP vol 2, 3rd edition, page 232.