# An analytical/semi-analytical analysis of the characteristics of the Expected Worst Loss in T realisations for Normal random variables

[Nematrian website page: ExpectedWorstLossAnalysis, © Nematrian 2015]

Suppose we have $X_i$ which are independent identically distributed Normal random variables, i.e. $X_i \sim N(\mu, \sigma^2)$. The worst loss in $T$ realisations is $X_{(1)} = min(X_1, \ldots, X_T)$.

What, in such circumstances is the expected worst loss (in $T$ realisations)?

Without much loss of generality, we can focus on $Y_i = (X_i - \mu)/\sigma$, i.e. unit Normal random variables, as $Y_{(1)} = \left(X_{(1)} - \mu\right)/\sigma$.

If $Y_{(1)} > u$ then all of the $Y_i$ must also satisfy $Y_i > u$. As the $Y_i$ are independent and identically distribution, this means that:

$$Pr\left(Y_{(1)} < u\right) = 1 - Pr\left(Y_{(1)} > u\right) = 1 - Pr(Y_i > u)^T = 1 - \left(1 - Pr(Y_i < u)\right)^T$$
$$= 1 - \left(1 - \int_{-\infty}^{u} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt\right)^T = 1 - \left(1 - N(u)\right)^T$$

where $N(z)$ is the cumulative distribution function of the unit Normal distribution.

So the probability density function of $Y_{(1)}$ is $p(u)$ where:

$$p(u) = \frac{d}{du} Pr\left(Y_{(1)} < u\right) = \frac{T\left(1 - N(u)\right)^{T-1}}{\sqrt{2\pi}} e^{-u^2/2}$$

The *expected* value of $Y_{(1)}$, i.e. the expected worst loss in $T$ realisations for a unit Normal distribution, is then $EWL(T) = E\left(Y_{(1)}\right)$, i.e.:

$$EWL(T) = \int_{-\infty}^{\infty} u.p(u)du = \int_{-\infty}^{\infty} u \frac{T\left(1 - \int_{-\infty}^{u} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt\right)^{T-1}}{\sqrt{2\pi}} e^{-u^2/2} du$$

Using a symbolic algebra engine we find that there are analytical formulae $EWL(T)$ for $T$ = 1, 2 or 3 but not thereafter:

| $T$ | $EWL(T)$ |
|---|---|
| 1 | $0$ |
| 2 | $-\frac{1}{\sqrt{\pi}}$ |
| 3 | $-\frac{3}{4}\frac{\sqrt{2}}{\pi}$ |

As $T$ increases $EWL(T)$ becomes more negative, but only relatively slowly:

| $T$ | $EWL(T)$ | $EWL(T) - EWL(T/2)$ |
|---|---|---|

| | | |
|---|---|---|
| 1 | 0 | |
| 2 | -0.5641895835 | -0.5641895835 |
| 4 | -1.029375373 | -0.4651857895 |
| 8 | -1.423600306 | -0.394224933 |
| 16 | -1.765991393 | -0.342391087 |
| 32 | -2.069668828 | -0.303677435 |
| 64 | -2.343733465 | -0.274064637 |
| 128 | -2.594597369 | -0.250863904 |
| 256 | -2.826863279 | -0.232265910 |

We may check that the values shown above are reasonable using a simulation approach, e.g. using VBA code in Microsoft Excel as per VBA code that can be used to check this analysis.

The worst loss can also be thought of as a specific quantile of a sample of Normal random variables. It might be viewed as the $1/(2T)$'th quantile (i.e. half way between 0 and $1/T$, given that the sample has $T$ observations) . We might therefore expect its expected value to be similar to the corresponding quantile point of the Normal distribution. However, the Normal probability density function is not normally approximately flat at the relevant quantile point (and is instead upward sloping), so the Expected Worst Loss in T realisations is normally somewhat above the $1/(2T)$'th Normal quantile point:

| $T$ | $EWL(T)$ | c.f. $N^{-1}\left(\frac{1}{2T}\right)$ |
|---|---|---|
| 1 | 0 | 0 |
| 2 | -0.5641895835 | -0.67449 |
| 4 | -1.029375373 | -1.15035 |
| 8 | -1.423600306 | -1.53412 |
| 16 | -1.765991393 | -1.86273 |
| 32 | -2.069668828 | -2.15387 |
| 64 | -2.343733465 | -2.41756 |
| 128 | -2.594597369 | -2.66007 |
| 256 | -2.826863279 | -2.88563 |