6.10 Inverse Distributions

The inverse of a function is another function that takes the output of the initial function as its input, and reproducing the input of the initial function as its output. The following are examples of functions and their inverses:

$$y = ax \rightarrow x = \frac{1}{a} y \tag{6.94}$$

or

$$y = x^2 \rightarrow x = \pm\sqrt{y} \tag{6.95}$$

or, in general[302]

$$y = f(x) \rightarrow x = f^{-1}(y) \tag{6.96}$$

The inverse is often described as mirror reflection in a 45^o line (try drawing one of the functions above by hand, and then its 45^o reflection as well).

An alternatively application of an inverse function[303] can exchange a sequence of paired data values *x's* and *y's*, say in plot, and then simply swapping the plotting inputs so that the *x's* become the *y's*, and *y's* and the *x's*.

In general, an inverse function may not exist, and "inverse function theory" is considered a complex area of study. One illustration of the complexity is that a function that has "nice properties" (see Chapter 3.11) may have an inverse function with "nasty properties" (e.g. multiple solutions, singularities, etc). For example, a horizontal line becomes a vertical line (i.e. a singularity). Another example is seen in Equation (6.95), which exhibits "uniqueness" difficulty since the "full solution" of the inverse problem produces both a "+" and "-" function simultaneously[304], and so it has multiple solutions for $y>0$ [305].

---

[302] Remember that although a superscript –1 often means "division", in the general context of inverse functions this superscript means "inverse". For example, $\text{Sin}^{-1}(x)$ is not the same thing as $\text{Sin}(x)^{-1}$.

[303] If the data are discrete data (as opposed to points generated by a continuous function), then the word "function" is replaced with the word "mapping".

[304] Recall that the square root of any function must be considered to have $\pm$ solutions since squaring either the "+" or "-" of any number will produce the original function.

[305] This "Root2 shape" will arise frequently in later Chapters, where it is shown that under different conditions the "uniqueness issue" is manageable.

In the context of distributions in trading, the "usual" calculation is to find a cumulative probability of an event (i.e. for an "input"). That is, find the quantile, given some "level" or "input" (such as strike price). Thus, the inverse problem is to find the "level" or "input", given a quantile or cumulative probability. VaR is exactly this type of inverse distribution problem.

As such, the inverse (distribution) problem requires finding the inverse distribution function (usually the inverse cumulative distribution function). There are two possible approaches: analytical inverse functions or numerical methods based solutions and approximations.

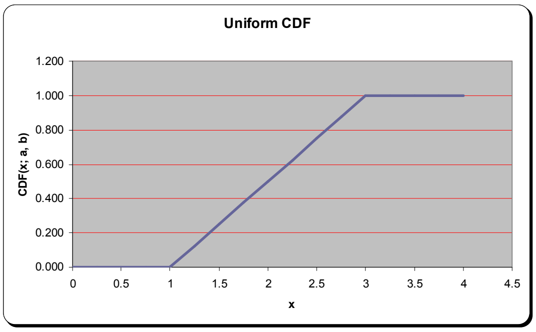
In this discussion, inverse distributions are considered only for continuous distributions. Analytical solutions may be expected to be difficult if not impossible, given the complex forms for some of the distributions in Section 11.3.2.3. As an illustration, though, consider the CDF for the Uniform distribution as:

$$\mathrm{CDF}_{\mathrm{Uniform}}\left[x \middle| a,b\right] = y = \begin{cases} 0 & x < a \\ \dfrac{x-a}{b-a} & , x \in \left[a,b\right] \\ 1 & x > b \end{cases} \tag{6.97}$$

The inverse of this function is:

$$\mathrm{CDF}^{-1}_{\mathrm{Uniform}}\left[y \middle| a,b\right] = x = \begin{cases} ? & x < a \\ y\left(b-a\right)+a & , x \in \left[a,b\right] \\ ? & x > b \end{cases} \tag{6.98}$$

Notice that even for what appears to be a relatively simple shape (i.e. one involving straight lines), the inverse is a bit messy since the horizontal components of the original CDF are now singular (vertical) lines, with undefined values. The CDF and CDF^{-1} are illustrated in Figure 6.10 – 1.



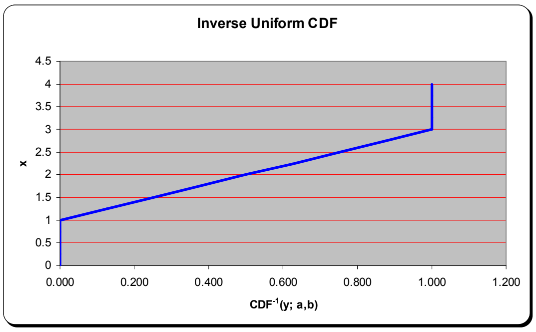


Figure 6.10 – 1. a) The CDF for the Uniform distribution, and b) the inverse CDF (i.e. CDF^{-1}) for the Uniform distribution.

Therefore, if Figure 6.10 – 1 a) depicts, say, the P&L (cumulative) distribution of a portfolio in preparation for VaR calculation, then simply specifying the VaR confidence level (say 5%) requires only to look along the horizontal axis of Figure 6.10 – 1 b) to the 0.05 level and reading the function value from the vertical axis. In this case, the 5%ntile corresponds to an input value of 1.1, which in a VaR threshold context would be interpreted as "the P&L should be worse than 1.1 only 5% of the time".

Numerical methods and quasi-closed form approximations (such as those discussed in Section 6.11) are likely to be the most common implementations in practice due to the intractable or cumbersome nature of the analytical solutions for many distributions. Perhaps surprisingly, in some cases the inverse may actually exist even when the target CDF is non-intergrable. However, to know when this is true may require input from your quant group. In many cases, numerical methods can be used successfully even when there is little chance of coming to grips with the analytical formulations. Though, and as can be seen above, the inverse functions will likely require much "fiddling" to ensure that the numerical implementation are not only correct, but that the limiting/truncated portions are handled sensibly.

Example of these types of numerical implementation are provided in Chapter 16, and [3.a] and [3.f].

For example, analytical inversion of the Normal CDF:

$$\mathrm{CDF}_{\mathrm{Normal}}\left[x\middle|\mu_x,\sigma_x\right]=\int_{-\infty}^{x}\frac{1}{\sigma_x\sqrt{2\pi}}\;e^{-\frac{1}{2}\left(\frac{x-\mu_x}{\sigma_x}\right)^2}dx \tag{6.99}$$

may not be too practical, and is much too daunting in any case. Rather, the inversion of a quasi-closed form fitted approximation of the form:

$$\begin{aligned}&\mathrm{CDF}_{\mathrm{Normal}}\left[x\middle|0,1\right]\cong\sum_{i=1}^{N}\beta_i\left(\frac{1}{1+\alpha|x|}\right)^i,\ \textit{when } x\geq 0\\&\textit{and}\\&\mathrm{CDF}_{\mathrm{Normal}}\left[x\middle|0,1\right]\cong 1-\mathrm{CumProb}_{\mathrm{Normal}}\left[|x|\middle|0,1\right],\ \textit{when } x<0\end{aligned} \tag{6.100}$$

is the likely real world approach, even if it requires "root finding numerical inversion" (see Chapter 10).

Additionally, some numerical methods lend themselves more easily to "effective" inverse calculations. For example, using a Rank based method to evaluate CDFs with histograms obviates the need for an explicit inversion. Then, however, the results are restricted to the "accuracy" of the Rank based estimation (as opposed to, say, a cubic spline fit, with inversion of the spline or root finding quadrature on the spline).