ArŶmα Labs
Decoding the past, Encoding the future

Robyn Under The Hood - Adstock Transformation

In this article, we will deep dive into what is Weibull Distribution and how it is applied to adstock media variables in MMM.

What is Weibull Distribution?

Weibull distribution is a continuous distribution which has the ability to assume the characteristics of many different types of distributions. It is most used in Reliability Engineering to understand time to failure for different equipment. It can also model skewed data (both left skewed and right skewed).

Weibull Distribution is defined by two versions:

1. Three parameter distribution: The three-parameter distribution has three parameters: shape, scale and location. Here, ϒ is the shape parameter, α is the scale parameter and μ is the location parameter. (Refer to Equation 1).

2. Two parameter distribution: The two-parameter distribution has 2 parameters: shape and scale and the location parameter, μ = 0. The shape parameter controls the shape of the decay curve. The smaller the value, the more L-shape the curve is and the larger the value, the more S-shape the curve is. The scale parameter controls the inflexion point of the decay curve. Basically, it controls how far stretched the distribution is.

Probability Density Function (PDF) for three – parameter distribution

$$f(x) = \frac{\gamma}{\alpha}\left(\frac{x-\mu}{\alpha}\right)^{(\gamma-1)} \exp\left(-((x-\mu)/\alpha)^{\gamma}\right) \qquad x \geq \mu; \gamma, \alpha > 0$$

Equation 1

PDF for two-parameter function would have μ = 0.

Similarly, CDF function for two-parameter distribution is:

$$F(x) = 1 - \exp(-(x/\alpha)^{\gamma}), \text{ for } x >= 0$$

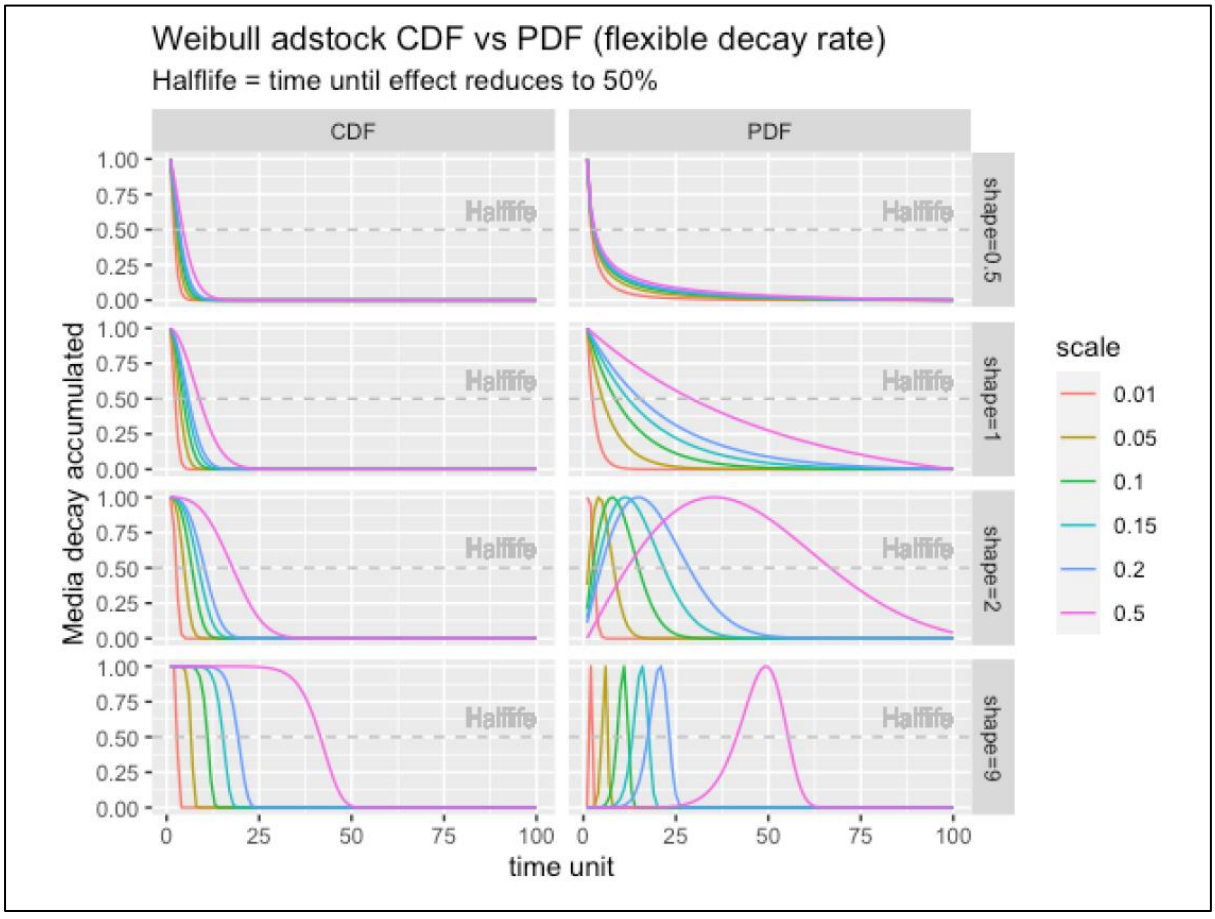
Equation 2

In case you want a refresher on what is PDF, check this intuitive explanation by Aerin Kim. (Link: https://towardsdatascience.com/pdf-is-not-a-probability-5a4b8a5d9531)

The Weibull Distribution function used in Robyn MMM code, is a two – parameter distribution where shape and scale parameters are tuned.

Why is Weibull Distribution function used for adstocking?

The decay rate used with the conventional adstock method like Geometric adstock is fixed but, Weibull adstock has two parameters (shape and scale) which provide more flexibility in terms of how the carryover effect can be modelled. The below image depicts how changing shape and scale parameters affect the shape and inflexion point of the PDF and CDF curves.



Source: Facebook Robyn

Weibull PDF and CDF Adstocks

The Weibull adstock is computed using two different transformations.

1. Weibull CDF Adstock: The CDF Adstock is applied using the pweibull function in R. The pweibull function generates the CDF value for each input. Refer to Code Snippet 1

2. Weibull PDF Adstock: The PDF Adstock is applied using the dweibull function in R. The dweibull function generates corresponding values of Weibull density for each input. Refer to Code Snippet 1.

```r
adstock_weibull <- function(x, shape, scale, windlen = length(x), type = "CDF") {
  check_opts(toupper(type), c("CDF", "PDF"))
  x_bin <- 1:windlen
  scaleTrans <- round(quantile(1:windlen, scale), 0)
  if (shape == 0) {
    thetaVecCum <- thetaVec <- rep(0, windlen)
  } else {
    if ("CDF" %in% toupper(type)) {
      thetaVec <- c(1, 1 - pweibull(head(x_bin, -1), shape = shape, scale = scaleTrans)) # plot(thetaVec)
      thetaVecCum <- cumprod(thetaVec) # plot(thetaVecCum)
    } else if ("PDF" %in% toupper(type)) {
      thetaVecCum <- .normalize(dweibull(x_bin, shape = shape, scale = scaleTrans)) # plot(thetaVecCum)
    }
  }
  x_decayed <- mapply(function(x_val, x_pos) {
    x.vec <- c(rep(0, x_pos - 1), rep(x_val, windlen - x_pos + 1))
    thetaVecCumLag <- lag(thetaVecCum, x_pos - 1, default = 0)
    x.prod <- x.vec * thetaVecCumLag
    return(x.prod)
  }, x_val = x, x_pos = x_bin)
  x_decayed <- rowSums(x_decayed)
  return(list(x = x, x_decayed = x_decayed, thetaVecCum = thetaVecCum))
}

.normalize <- function(x) {
  if (diff(range(x)) == 0) {
    return(c(1, rep(0, length(x) - 1)))
  } else {
    return((x - min(x)) / (max(x) - min(x)))
  }
}
```

Distribution Function

Density Function

Code Snippet 1

The user has to specify whether to use CDF or PDF while running the “adstock_weibull” function and finally the transformed series x_decayed is generated.

Resources:

1. https://github.com/facebookexperimental/Robyn/blob/main/R/R/transformation.R
2. https://facebookexperimental.github.io/Robyn/docs/features/
3. https://www.itl.nist.gov/div898/handbook/eda/section3/eda3668.htm

4. https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/Weibull
5. https://statisticsglobe.com/weibull-distribution-in-r-dweibull-pweibull-qweibull-rweibull
6. https://statisticsbyjim.com/probability/weibull-distribution/