THOUGHTS FOR THE QUASI-WEIBULL MODEL

STEVEN J. MILLER

1. Random Variable Formulation

We are trying to model runs scored and allowed from a quasi-Weibull, where there is a point mass at the origin and then a Weibull with parameters $\alpha, \beta = 1/2$ and γ. We assume the probability we get 0 is p (i.e., the probability we draw from the delta function is p) and thus the probability we draw from the Weibull is $1-p$.

Let X_p be the random variable which is 1 with probability p and 0 with probability $1-p$. Let X_δ be the random variable which is 0 with probability 1, and let X_w be the random variable which has the Weibull distribution with parameters α, β and γ. Then

$$X \ = \ X_p X_\delta + (1 - X_p) X_w$$

is the random variable denoting the runs scored by our team.

Note that the variables X_p, X_δ and X_w are independent but clearly X_p and $1 - X_p$ are not independent.

2. Mean

We first calculate the mean value. Note the expected value of the product of independent random variables is the product of the expected values. We have

$$\begin{aligned} \mathbb{E}[X] &= \mathbb{E}[X_p X_\delta + (1 - X_p) X_w] \\ &= \mathbb{E}[X_p]\mathbb{E}[X_\delta] + \mathbb{E}[(1 - X_p)]\mathbb{E}[X_w] \\ &= p \cdot 0 + (1-p)\mathbb{E}[X_w] \\ &= (1-p)\mu_w, \end{aligned} \tag{2.1}$$

where μ_w is the mean of a Weibull with parameters α, β, γ.

3. Variance

To compute the variance, we use the formula $\mathrm{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$. We already know $\mathbb{E}[X] = \mu_w$, so it suffices to compute $\mathbb{E}[X^2]$. We have

$$\begin{aligned} \mathbb{E}[X_p^2] &= 1^2 \cdot p + 0^2 \cdot (1-p) \ = \ p \\ \mathbb{E}[(1-X_p)^2] &= 0^2 \cdot p + 1^2 \cdot (1-p) \ = \ 1-p \\ \mathbb{E}[X_p(1-X_p)] &= 1(1-1) \cdot p + 0(1-0) \cdot (1-p) \ = \ 0 \\ \mathbb{E}[X_\delta] &= 1 \\ \mathbb{E}[X_\delta^2] &= 1 \\ \mathbb{E}[X_w] &= \mu_w \\ \mathbb{E}[X_w^2] &= \mathrm{Var}(X_w) + \mu_w^2. \end{aligned} \tag{3.1}$$

We now compute the second moment of X.

$$\begin{aligned}
\mathbb{E}[X^2] &= \mathbb{E}\left[(X_pX_\delta + (1-X_p)X_w)^2\right] \\
&= \mathbb{E}\left[X_p^2X_\delta^2 + 2X_p(1-X_p)X_\delta X_w + (1-X_p)^2X_w^2\right] \\
&= \mathbb{E}\left[X_p^2X_\delta^2\right] + 2\mathbb{E}\left[X_p(1-X_p)X_\delta X_w\right] + \mathbb{E}\left[(1-X_p)^2X_w^2\right] \\
&= \mathbb{E}\left[X_p^2\right]\mathbb{E}\left[X_\delta^2\right] + 2\mathbb{E}\left[X_p(1-X_p)\right]\mathbb{E}\left[X_\delta\right]\mathbb{E}\left[X_w\right] + \mathbb{E}\left[(1-X_p)^2\right]\mathbb{E}\left[X_w^2\right] \\
&= p\cdot 0 + 2\cdot 0\cdot 1\cdot \mu_w + (1-p)\cdot\left(\mathrm{Var}(X_w)+\mu_w^2\right) \\
&= (1-p)\cdot\left(\mathrm{Var}(X_w)+\mu_w^2\right), \qquad (3.2)
\end{aligned}$$

which implies

$$\begin{aligned}
\mathrm{Var}(X) &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 \\
&= (1-p)\cdot\left(\mathrm{Var}(X_w)+\mu_w^2\right) - ((1-p)\mu_w)^2 \\
&= (1-p)\mathrm{Var}(X_w) + (1-p)(1-(1-p))\mu_w^2 \\
&= (1-p)\mathrm{Var}(X_w) + p(1-p)\mu_w^2 \qquad (3.3)
\end{aligned}$$

Does this formula make sense? Let's consider what happens if μ_w becomes large, say 100. In this case, the point mass at the origin should greatly increase the variance (i.e., the variance of X should greatly exceed that of X_w). The reason is that we now have a lot of mass far away from the Weibull part. Note our formula does have the variance increase as μ_w moves away from 0 (in either direction). In the special case when $\mu_w = 0$ then the effect of placing p percent of the mass at the origin should make the variance smaller, and we see that our formula captures that as well.

The best way to be sure, of course, is to simulate data, drawing from the delta mass at the origin p percent of the time and from our Weibull $1-p$ percent of the time.

Department of Mathematics and Statistics, Williams College
E-mail address: sjm1@williams.edu