Gamma LRT

This document describes the steps to obtain the tail function for the importance sampling weights in Gamma distribution example. We use an Exponential distribution as the proposal dist with pdf $g(x)$, and our target distribution is Gamma with pdf $f(x)$.

Steps are as follows (I think!) This looks really good Payman. I just added a couple notes in blue.

Step 1. Set the target distribution, $f(x)$ as a Gamma distribution with α as shape and β as scale:

$$f(x; \alpha, \beta) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{\beta x} \tag{1}$$

Step 2. Set the proposal distribution, $g(x)$, as an Exponential dist with scale parameter of one:

$$g(x) = e^{-x} \tag{2}$$

More generally, we can use any scale parameter for the exponential. For the purposes of this discussion, we just need the scale parameter to be fixed. I think it's pretty reasonable to keep $\lambda = 1$ for now to avoid cluttering our formulas, although eventually we'll need to do the general case.

Step 3. The pdf of weights is $w(x; \alpha, \beta) = \frac{f(x; \alpha, \beta)}{g(x)}$. This actually isn't quite the pdf of the weights. Instead, the Xs are iid from g, and the weights are obtained by evaluating w at the simulated Xs. The distribution on the weights is then induced by the distribution on X and the functional relationship between X and W. If the words "pushforward measure" help clarify things, that's exactly what's happening here. If not, W is a transformation of X. In principle, we can get the pdf of W by computing the Jacobian of the (inverse) transformation (i.e. the derivative of the inverse of the likelihood ratio function). Personally, I prefer to work directly with the CDF or survival function; for one, I think it's easier, and for two, we need the survival function for step 4 anyway.

Step 4. Use the pdf in step 3 and compute its tail function as described in Pickands paper:

$$P(x) = \frac{F(X \geq u + x)}{F(X \geq u)} \tag{3}$$

where $F(.)$ is the CDF of weights as defined in step 3.

Step 5. Find Q_3, and median as described in the paper. Compute the c and a.

1 William's Derivation

I'm going to mostly stick to the above notation, but I would like to use the general exponential distribution until I find myself needing to set $\lambda = 1$. I will also suppress dependence on parameters in the various densities to avoid extra typing. Thus, $g(x) = \exp(-x/\lambda)/\lambda$.

At a high-level, my plan is to compute the survival function of the weights, as well as their density and the derivative of their density. I can then use Equation (1.1.30) on page 15 of de Haan and Ferreira (2006) to (hopefully) compute the tail index. There are enough steps here that it's going to take a few days' work. One step at a time.

To start, we note that the weight, W, can be written as a random variable in the following way.

$$W(X) = \frac{f(X)}{g(X)} \tag{4}$$

$$= \frac{\lambda}{\beta^\alpha \Gamma(\alpha)} X^{\alpha-1} \exp\left[-X\left(\frac{1}{\beta} + \frac{1}{\lambda}\right)\right] \tag{5}$$

$$=: cX^{\alpha-1} \exp(-X/\gamma) \tag{6}$$

where $c = \lambda/\beta^\alpha \Gamma(\alpha)$, and $\gamma = (1/\beta + 1/\lambda)^{-1}$. I doubt that it will every be relevant, but I think it's cute that γ is a harmonic mean.

We can now work out the survival function of W. We write $\mathscr{W}$ for the inconveniently named Lambert-W function, which is the inverse function of $x \mapsto xe^x$. I.e. $\mathscr{W}$ satisfies $\mathscr{W}(x) \exp[\mathscr{W}(x)] = x$.

$$\mathbb{P}(W > w) = \mathbb{P}\left(X^{\alpha-1} e^{-X/\gamma} > \frac{w}{c}\right) \tag{7}$$

$$= \mathbb{P}\left(X > -(\alpha-1)\gamma \cdot \mathscr{W}\left[-\frac{(w/c)^{1/(\alpha-1)}}{(\alpha-1)\gamma}\right]\right) \tag{8}$$

$$= \exp\left(\frac{(\alpha-1)\gamma}{\lambda} \cdot \mathscr{W}\left[-\frac{(w/c)^{1/(\alpha-1)}}{(\alpha-1)\gamma}\right]\right) \tag{9}$$

where we pass from (7) to (8) with the help of Wolfram Alpha.

1.1 PDF

Next, we need the PDF of W, and later its derivative. I'm going to find it easier to do this in stages. Let $\bar{\varphi}(w) = \mathbb{P}(W > w)$ and η be the argument to $\mathscr{W}$ in (9). The (negative) PDF of w is obtained by differentiating $\bar{\varphi}$. Note that the

derivative of $\mathscr{W}$ can be obtained implicitly from its[1] defining equation

$$\nabla_w \bar{\varphi}(w) = \frac{(\alpha - 1)\gamma}{\lambda} \bar{\varphi}(w) \nabla_w \mathscr{W}(\eta) \tag{10}$$

$$\nabla_w \mathscr{W}(\eta) = \frac{\mathscr{W}(\eta)}{\eta[1 + \mathscr{W}(\eta)]} \nabla_w \eta \tag{11}$$

$$\nabla_w \eta = -\frac{1}{c\gamma(\alpha - 1)^2} \left(\frac{w}{c}\right)^{\frac{1}{\alpha-1} - 1} = \frac{\eta}{\alpha - 1} \tag{12}$$

Putting this all together, and after a surprising amount of cancellation, I get

$$f_W(w) := -\nabla_w \bar{\varphi}(w) \tag{13}$$

$$= -\frac{\bar{\varphi}(w)}{\lambda w} \cdot \frac{\mathscr{W}(\eta)}{1 + \mathscr{W}(\eta)} \tag{14}$$

1.2 Derivative of 1 over Hazard Rate

We return now to Theorem 1.1.8 of de Haan and Ferreira (2006), specifically expression $(1.1.29)$[2]. We first compute the survival function of W over its density, wherein we get some nice cancellation.

$$\frac{\bar{\varphi}(w)}{f_W(w)} = -\frac{\lambda w[1 + \mathscr{W}(\eta)]}{\mathscr{W}(\eta)} \tag{15}$$

where $\mathscr{W}$ and η are defined in Section 1.1. We now need to differentiate the expression in 15. Fortunately, we can recycle identities (8) and (9). We first present an identity which helps with simplification, then proceed with computing the derivative.

$$\nabla_w \mathscr{W}(\eta) = \partial \mathscr{W}(\eta) \cdot \partial \eta \tag{16}$$

$$= \frac{\mathscr{W}(\eta)}{(\alpha - 1)[1 + \mathscr{W}(\eta)]} \tag{17}$$

$$\nabla_w \left[\frac{\bar{\varphi}(w)}{f_W(w)}\right] = -\left(\frac{\lambda \mathscr{W}(\eta)\nabla_w[w(1 + \mathscr{W}(\eta))] - \lambda w[1 + \mathscr{W}(\eta)]\nabla_w \mathscr{W}(\eta)}{\mathscr{W}(\eta)^2}\right) \tag{18}$$

$$= -\left(\frac{\lambda \mathscr{W}(\eta)[(1 + \mathscr{W}(\eta)) + w\mathscr{W}(\eta)/([\alpha - 1][1 + \mathscr{W}(\eta)])] - \lambda w\mathscr{W}(\eta)/(\alpha - 1)}{\mathscr{W}(\eta)^2}\right) \tag{19}$$

$$= \lambda \left[\frac{w}{(\alpha - 1)\mathscr{W}(\eta)(1 + \mathscr{W}(\eta))} - \frac{(1 + \mathscr{W}(\eta))}{\mathscr{W}(\eta)}\right] \tag{20}$$

---

[1]Usually, I try not to use "it" to refer to anything in academic writing (Tom drilled this into me during my master's thesis revisions). Here however, I think it's a pretty good fit and doing anything else would be very wordy.

[2]Note that, in survival analysis, the hazard rate of a distribution is defined as the density over the survival function. The quantity being differentiated in (1.1.29) before taking the limit is thus the reciprocal of the hazard rate of W

All that remains is to compute the limit of (20) as $w \uparrow \infty$. Note that $\eta \to -\infty$ as $w \to \infty$, so $\mathscr{W}(\eta) \to -\infty$ as well. Re-write (20) as $\lambda[A - B]$. We now explore the asymptotic behaviour of A and B.

First, it's not hard to show that $B \to 1$. However, A takes more work. Recall that $\mathscr{W}(\eta)\exp[\mathscr{W}(\eta)] = \eta$, so we can re-write A as

$$A = \frac{\eta}{\mathscr{W}(\eta)} \cdot \frac{w}{(\alpha - 1)\eta[1 + \mathscr{W}(\eta)]} \tag{21}$$

$$= \exp[\mathscr{W}(\eta)] \cdot \frac{w}{(\alpha - 1)\eta[1 + \mathscr{W}(\eta)]} \tag{22}$$

$$= \frac{1}{\alpha - 1} \cdot w \cdot \frac{\exp[\mathscr{W}(\eta)]}{\eta[1 + \mathscr{W}(\eta)]} \tag{23}$$

$$=: \frac{1}{\alpha - 1} \cdot w \cdot \xi(\eta) \tag{24}$$

Unfortunately, I don't have any a priori information about the relative growth rates of w and η (or any functions of η). I do know that $\xi(\eta) \to 0$ as $w \to \infty$, so we can re-write $(\alpha - 1)A = w/[1/\xi(\eta)]$ and apply l'Hospital's Rule.

$$\frac{w}{1/\xi(\eta)} \sim \frac{1}{\nabla_w[1/\xi(\eta)]} \tag{25}$$

$$\nabla_w \left[\frac{1}{\xi(\eta)} \right] = \nabla_w \left(\frac{\eta[1 + \mathscr{W}(\eta)]}{\exp[\mathscr{W}(\eta)]} \right) \tag{26}$$

$$= \nabla_\eta \left(\frac{\eta[1 + \mathscr{W}(\eta)]}{\exp[\mathscr{W}(\eta)]} \right) \cdot \nabla_w \eta \tag{27}$$

$$= \frac{\mathscr{W}(\eta)}{\alpha - 1} \cdot \frac{1 + 2\mathscr{W}(\eta)}{1 + \mathscr{W}(\eta)} \tag{28}$$

Where the last step comes from much expanding and simplifying. Finally, we get

$$A \sim \frac{1}{\alpha - 1} \left[\frac{\mathscr{W}(\eta)}{\alpha - 1} \cdot \frac{1 + 2\mathscr{W}(\eta)}{1 + \mathscr{W}(\eta)} \right]^{-1} \tag{29}$$

$$= \frac{1}{(\alpha - 1)^2} \frac{1 + \mathscr{W}(\eta)}{\mathscr{W}(\eta)[1 + 2\mathscr{W}(\eta)]} \tag{30}$$

$$\sim \frac{1}{2(\alpha - 1)^2} \cdot \frac{1}{\mathscr{W}(\eta)} \tag{31}$$

$$\to 0 \tag{32}$$

Putting everything together, our expression for the tail index is

$$\lim_{w \uparrow \infty} \nabla_w \left[\frac{\bar{\varphi}(w)}{f_W(w)} \right] = \lambda[A(w) - B(w)] \tag{33}$$

$$= \lambda[o(1) - 1 + o(1)] \tag{34}$$

$$\to -\lambda \tag{35}$$

4

1.3   Discussion

This final conclusion, while pleasantly simple, cannot possibly be correct. I can't believe that the tail index of W doesn't depend on the parameters of the target distribution. Frankly, an incorrect answer isn't terribly surprising here. It's a long derivation with lots of room for mistakes. If I have time between now and our next meeting, I'll spend some time reviewing my work. If not, I would like to go over it again with you Payman, both to explain my reasoning and to double-check my work.

References

de Haan, L. and Ferreira, A. (2006). *Extreme value theory: an introduction.* Springer.