An Actuary Walks into a War Room (and other jokes)

MPSA, 3-6 April

Miles D. Williams, Denison University
(williamsmd@denison.edu)

The Problem

- War deaths are an unruly thing to model.
- When Cunen, Hjort, and Nygård (2020) turned to an unconventional model called the inverse Burr to study war size, they were drawing from a deep well of actuarial science techniques to solve this problem.
- But they say little about what this means, and offer even less practical advice and an easy to use user interface for working with the model.

My Solution

- I created an R package (`{invburreg}`) that implements an inverse Burr regression in ways that are consistent with other modeling functions in R.
- I offer guidance on how to talk about model specification.
- And guidance on summarizing model results.

Background

- **What's different about an actuarial approach to modeling war size?**
- We political scientists tend to focus on expected outcomes — e.g., conditional averages.
- (From what I can tell) Actuaries take a whole-of-distribution approach to assessing risk.
- Actuaries need flexible models that can accommodate potentially unruly data.

Why does this matter for war?

- There's arowing literature testing the *decline of war thesis* (Braumoeller 2019; Cederman 2003; Cederman, Warren, and Sornette 2011; Cirillo and Taleb 2016; Clauset 2017, 2018; Spagat and Weezel 2020; Spagat, Johnson, and Weezel 2018).
- The decline of war debate is concerned with whether wars are becoming less frequent over time and less deadly.
- When it comes to testing the second part of this argument, there's a problem.
- **Wars are prone to power-law behavior in the extreme tail of the distribution.**

The typical approach

Most applied research uses the classic power-law model to test whether there has been a change in the expected size of war over time:

$$\text{Pr}(X > x) \propto 1/x^{\alpha} \quad : \quad \alpha > 0, \forall \text{ large } x$$

Because of practical limitations, the power-law usually must be fit to outcomes greater than some threshold identified by way of GOF tests.

The typical approach

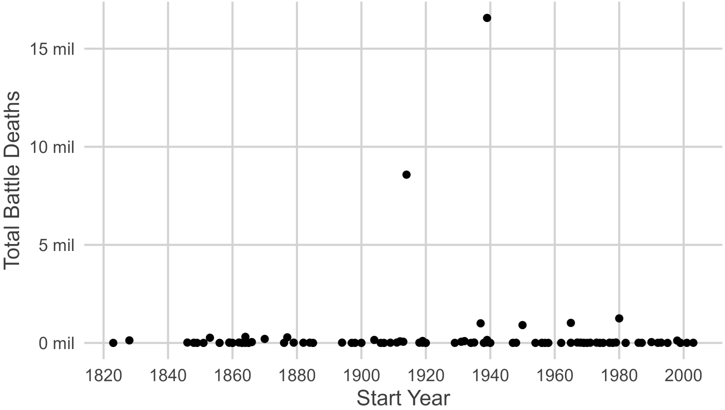

Figure 1: The total number of battle deaths by war in the Correlates of War conflict series. Values shown by war start year.

The typical approach

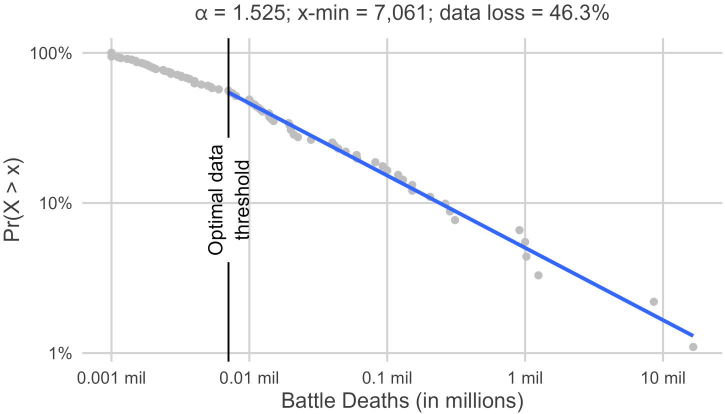

Figure 2: The optimal power-law fit for the battle death series. Values are shown on the log-scale.

Limitations

- Data truncation is necessary but comes at a cost to precision.
- Up to half of wars in the data remain unexplained.
- The classic power-law isn't amenable to covariates in a regression framework.

The solution

- Cunen, Hjort, and Nygård (2020) proposed using the inverse Burr as an alternative to the classic power-law for modeling war size.
- The inverse Burr is a **three parameter** distribution that can accommodate thick-tailed data while being flexible enough to capture an interior mode (unlike the classic power-law).
- It also can be modified to incorporate covariates into its estimation (also unlike the classic power-law).

The solution

The inverse Burr specifies that the probability of some event bigger than size x is:

$$\mathsf{Pr}(X > x) = 1 - \left[\frac{(x/\mu)^\theta}{1 + (x/\mu)^\theta}\right]^\alpha$$

α and θ are shape parameters that primarily (but not exclusively) influence the left and right-hand sides of the distribution, and μ is a scale parameter that influences the mode.

The solution

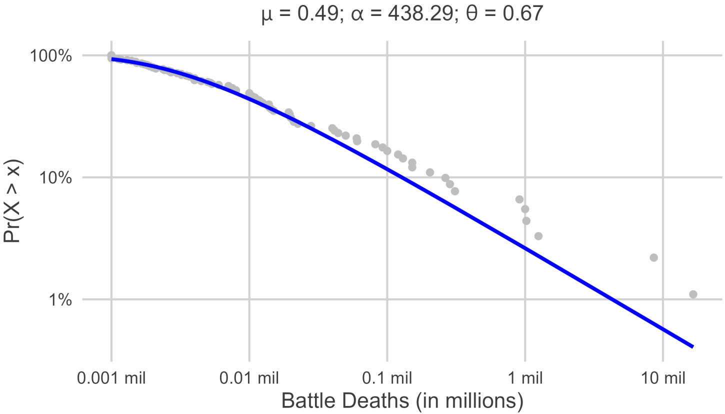

Figure 3: The optimal inberse Burr fit for the battle death series. Values are shown on the log-scale.

The solution

The simple inverse Burr does under-predict the most extreme wars, but Cunen, Hjort, and Nygård (2020) showed how the inverse Burr parameters can be fit using covariates (which also improves the overall fit of the model).

$$\mu_{L,i} = \mu_{L,0} \mathsf{exp}(\beta_L w_i) \quad \text{and} \quad \mu_{R,i} = \mu_{R,0} \mathsf{exp}(\beta_R w_i)$$

with constant α.

Making the solution accessible

- Cunen, Hjort, and Nygård (2020) provide replication materials and a short summary of their approach.
- But their goal was not to provide guidance on how to use this kind of model — it was to test the decline of war thesis.
- They offer little general advice about the model and how to specify an inverse Burr regression.
- They provide their R code, but it isn't exactly tidy and easy to use.
- They don't address appropriate ways to summarize model results.

Making the solution accessible

1. Model specification – both formally and using an R package I created called `{invburreg}`.
2. Model estimation – by way of maximum likelihood using bootstrapping for inference.
3. Model evaluation – simulating outcomes conditional model covariates.

Model specification

The inverse Burr density function with conditional model parameters:

$$\text{Burr}^{-1}(x_i; \mu_i, \alpha_i, \theta_i) = \frac{\alpha_i \theta_i (x_i/\mu_i)^{\alpha_i \theta_i}}{x_i[1 + (x_i/\mu_i)^{\theta_i}]^{\alpha_i + 1}}$$

where we have:

$$\begin{aligned}
\mu_i &= \exp(\mathbf{W}'\beta) \\
\alpha_i &= \exp(\mathbf{Y}'\delta) \\
\theta_i &= \exp(\mathbf{Z}'\gamma)
\end{aligned}$$

Model specification

Consider an applied example:

$$\mu_i = \text{exp}[\beta_0 + \beta_1 \text{post}_i + \beta_2 \text{polity}_i + \beta_3 \text{log}(\text{pop}_i)]$$
$$\alpha_i = \text{exp}[\delta_0 + \delta_1 \text{post}_i + \delta_2 \text{polity}_i + \delta_3 \text{log}(\text{pop}_i)]$$
$$\theta_i = \text{exp}[\gamma_0 + \gamma_1 \text{post}_i + \gamma_2 \text{polity}_i + \gamma_3 \text{log}(\text{pop}_i)]$$

The Outcome is battle-related deaths for 95 wars in the CoW conflict series (Sarkees and Wayman 2010).

The explanatory factors are a post-1950 dummy, average belligerent polity (Marshall, Gurr, and Jaggers 2017), and the log of pooled belligerent populations (Singer, Bremer, and Stuckey 1972).

Model specification

Using the `{invburreg}` package you can specify this model by writing:

```r
## open the package and access the wars example data
library(invburreg)
data("wars")

## fit an inverse Burr model to data
model_fit <- ibm(
  outcome = fat,
  mu = ~ post1950 + dem + pop,
  alpha = ~ post1950 + dem + pop,
  theta = ~ post1950 + dem + pop,
  data = wars
)
```

Estimation

The recommended approach to estimating the parameters of the inverse Burr model is MLE (Dey, Al-Zahrani, and Basloom 2017):

$$L = \sum_{i=1}^{n} \log[\mathsf{Burr}^{-1}(x_i; \hat{\mu}_i, \hat{\alpha}_i, \hat{\theta}_i)]$$

Cunen, Hjort, and Nygård (2020) use the Hessian for standard errors, but this doesn't always work in practice. `{invburreg}` relies on bootstrapping instead.

Evaluating results

- There are two useful ways to summarize results.
- The first is by way of a regression table (these are obligatory in all political science studies after all).
- The second is through simulating outcomes based on conditional inverse Burr parameters.

Evaluating results

Table 1: Inverse Burr regression estimates with bootstrapped standard errors.

	Baseline	Covariates	Baseline	Covariates	Baseline	Covariates
	log(α)		log(μ)		log(θ)	
Constant	6.083.	0.326	-0.708	0.224	-0.403***	-0.776**
	(3.693)	(0.499)	(5.218)	(0.365)	(0.085)	(0.28)
Post-1950		-0.449		-0.054		0.053
		(0.499)		(0.282)		(0.136)
Polity (Avg.)		0.266*		-0.007		0.042
		(0.116)		(0.139)		(0.026)
Population (log)		0.511***		-0.093		0.027
		(0.148)		(0.208)		(0.028)
	N = 95	N = 95	N = 95	N = 95	N = 95	N = 95

Note: .p < 0.1; *p < 0.05; **p < 0.01; ***p < 0.001

Evaluating results

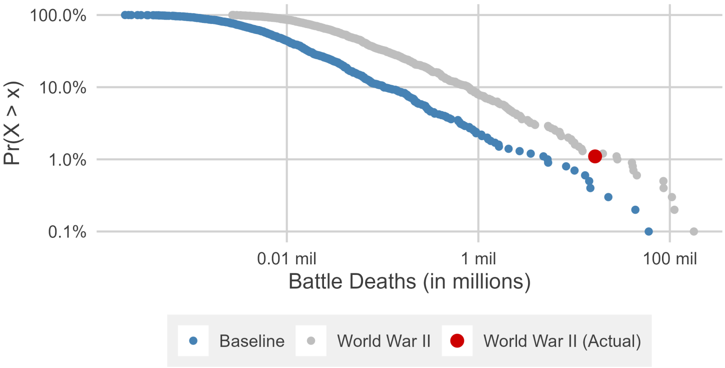

Figure 7: Simulated inverse Burr distributions based on the basline model and the model fit with covariates. In the covariate model, variables are set to World War II values. The actual Pr(X > x) for World War II is highlighted with a red point. Values are on the log-log scale

Evaluating results

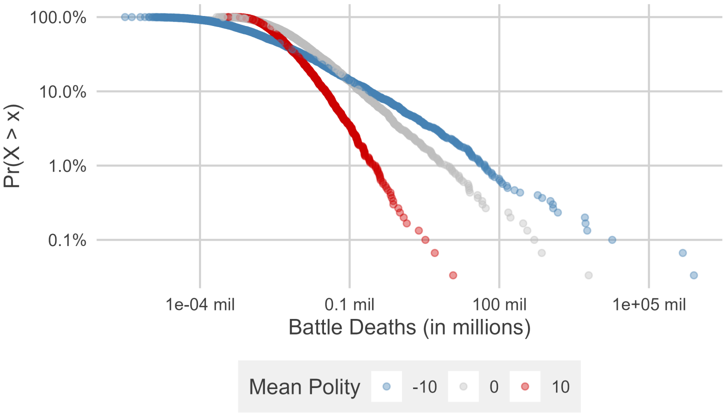

Figure 8: The conditional impact of average polity scores among belligerents on the CDF of war fatalities.

Warning!

Table 2: Inverse Burr regression estimates with bootstrapped standard errors. Unadjusted fatality totals (1) compared with adjusted fatalities (2) as the outcome.

	(1)	(2)	(1)	(2)	(1)	(2)
	log(α)		log(μ)		log(θ)	
Constant	0.326	-0.638	0.224	1.014*	-0.776**	-0.697*
	(0.499)	(0.476)	(0.365)	(0.482)	(0.28)	(0.32)
Post-1950	-0.449	-0.345	-0.054	0.066	0.053	0.066
	(0.499)	(0.456)	(0.282)	(0.641)	(0.136)	(0.196)
Polity (Avg.)	0.266*	0.032	-0.007	-0.189	0.042	0.007
	(0.116)	(0.07)	(0.139)	(0.155)	(0.026)	(0.032)
Population (log)	0.511***	0.068	-0.093	0.658**	0.027	0.011
	(0.148)	(0.099)	(0.208)	(0.237)	(0.028)	(0.033)
	N = 95	N = 95	N = 95	N = 95	N = 95	N = 95

Note: .p < 0.1; *p < 0.05; **p < 0.01; ***p < 0.001

Wrapping up

- The inverse Burr is a flexible tool used by actuaries to model risk in extreme-tailed data while modeling the whole of the distribution.
- It can be a great resource for conflict scholars studying correlates of war deaths.
- But a ready-to-use statistical package and guidance on best-practices is missing in the political science literature.
- This study is a first pass at making this a reality.

References

Braumoeller, Bear F. 2019. *Only the Dead: The Persistence of War in the Modern Age*. New York: Oxford University Press.

Cederman, Lars-Erik. 2003. "Modeling the Size of Wars: From Billiard Balls to Sandpiles." *American Political Science Review* 97 (1): 135–59.

Cederman, Lars-Erik, T. Camber Warren, and Didier Sornette. 2011. "Testing Clausewitz: Nationalism, Mass Mobilization, and the Severity of War." *International Organization* 65 (4): 605–38.

Cirillo, Pasquale, and Nassim Nicholas Taleb. 2016. "On the Statistical Properties and Tail Risk of Violent Conflicts." *Physica A: Statistical Mechanics and Its Applications* 452: 29–45.

Clauset, Aaron. 2017. "The Enduring Threat of a Large Interstate War." Technical report. One Earth Foundation.

———. 2018. "Trends and Fluctuations in the Severity of Interstate Wars." *Science Advances* 4 (2): eaao3580.

Cunen, Céline, Nils Lid Hjort, and Håvard Mokleiv Nygård. 2020. "Statistical Sightings of Better Angels: Analysing the Distribution of Battle-Deaths in Interstate Conflict over Time." *Journal of Peace Research* 57 (2): 221–34.

Dey, Sanku, Bander Al-Zahrani, and Samerah Basloom. 2017. "Dagum Distribution: Properties and Different Methods of Estimation." *International Journal of Statistics and Probability* 6 (2): 74–92.

Marshall, Monty G., Ted Robert Gurr, and Keith Jaggers. 2017. "Polity IV Project: Political Regime Characteristics and Transitions, 1800-2016."

Sarkees, Meredith Reid, and Frank Wayman. 2010. *Resort to War: 1816 - 2007*. Washington DC: CQ Press.

Singer, J. David, Stuart A. Bremer, and John Stuckey. 1972. "Capability Distribution, Uncertainty, and Major Power War, 1820-1965." In *Peace, War and Numbers*, edited by Bruce Russett. Beverly Hills, CA: Sage Publications, Inc.

Spagat, Michael, Neil F Johnson, and Stijn van Weezel. 2018. "Fundamental Patterns and Predictions of Event Size Distributions in Modern Wars and Terrorist Campaigns." *PloS One* 13 (10): e0204639.

Spagat, Michael, and Stijn van Weezel. 2020. "The Decline of War Since 1950: New Evidence." In *Lewis Fry Richardson: His Intellectual Legacy and Influence in the Social Sciences*, edited by Nils Peter Gleditsch, 129–42. Springer.