

Political Science 150C/350C
Spring 2007
Problem Set 3

1. The Poisson probability mass function (pmf) is

$$f(y; \lambda) = \frac{\lambda^y e^{-\lambda}}{y!}, \quad y = 0, 1, 2, \dots$$

Show that this pmf sums to one over its support, $y \in \{0, 1, 2, \dots\}$.

2. Show that if $Y \sim \text{Poisson}(\lambda)$, then $E(Y) = \lambda$. Hint: $E(y) = \sum_{y=0}^{\infty} y \cdot \Pr(Y = y)$.
3. Jim Fearon has kindly supplied us with cross-sectional data from $n=160$ countries, observed either in 1980 or in the first of independence (if that is after 1980). The [data](#) are available from my web site in Stata format; read them into R using the commands

```
library(foreign)
fearon <- read.dta(file="jackmanzinb.dta")
```

For each country we have a dependent variable, `totwars`, a count of the number of civil wars in country i from 1945-99 (an integer ranging from zero to six). The chief predictors (suggested to me by Fearon) are

- `lgdpenl1`, log of GDP per capita (based on the Penn World Table), lagged one year.
- `lpopl1`, log of population, lagged one year.
- `lmtnest`, log of percent of the country considered mountainous
- `Oil`, an indicator coded 1 if the country generates more 1/3 of its export revenues from fuels, and coded zero otherwise.

Use these predictors to model `totwars`. Write a **short** research note summarizing your findings. Do not be overly concerned with the theoretical setup. Focus tightly on (1) whether the Poisson is a suitable model for these data; (2) the nature of excess zeros (i.e., try fitting a logit model of the zeros against any other outcome); (3) the relative magnitudes of impact of predictors; (4) the possibility that outliers are influencing the results.

4. Hegre et al. (2001, *APSR* V95(1):33-48) analyze time until civil war onset with a Cox proportional hazards model. The data used in their analysis are available from [my web site](#) in Stata format (in the form received from Hegre et al.). The data are 11,560 observations at the country-spell level, meaning

that each observation is labelled by the start (`start`) and end (`stop`) dates of an observational period, the status of the observation (`status`), and a country label (`ss.numbe`, which I have kindly turned into an actual label for you, `name`). Additional details:

- `start`, days since Jan 1, 1816 (see the function `julian` for turning that into something human-readable, if you so desire).
- `stop`, days since Jan 1, 1816
- `status`, 0 if no event, 1 if an event (civil war outbreak).

The predictors in the Hegre et al model are:

- `prc`: temporal proximity of regime change
- `demo`: Polity score; note that Hegre et al. also include the square of this variable, since they are interested in a “U-shaped” pattern between levels of democracy and civil war duration
- `pcw`: temporal proximity of previous civil war; non-linear in calendar time, see Hegre et al. (2001), p36.
- `pi`: temporal proximity of independence; also non-linear in calendar time.
- `interwar`: dummy variable, 1 if international war in country, 0 otherwise.
- `neighbwa`: dummy variable, 1 if there is a civil war in a neighboring country.
- `ln.energ`: natural log of energy consumption per capita, measured in coal-ton equivalents; Hegre et al. also include the square of this variable, `energsq`.
- `ethnic.h`: a measure of ethnic heterogeneity.

Additional detail appears in the Hegre et al. article.

- (a) Replicate the findings in Hegre et al. Table 1 (p38) of their article; n.b., put a 95% confidence interval around the estimated half-lives. Also provide graphical displays of the Kaplan-Meier survival functions.
- (b) Replicate the post-1945 findings in Hegre et al Table 2 (p39) of their article. Note that the standard errors are “robust” standard errors produced by clustering on country; see the `robust` and `cluster` options in the `coxph` command. Does this “clustering” to produce the standard errors effect the results in any way?
- (c) Produce a graph showing how well the model fits the data. Comment on how well you think the model fits the data. Consult the chapter from *MASS* that is on my web site if you need suggestions.

- (d) Try to replicate something akin to Figure 1 Hegre et al; i.e., a graphical demonstration of the joint contribution of the democracy and years since regime change variables on the relative risk of civil war. Work carefully through the footnotes etc in the Hegre et al. article.

Due in class, Monday, May 21, 2007.