

PLSC 504 – Autumn 2017

Panel Models for Censored And Count Responses

November 2, 2017

“Lower” censored Y :

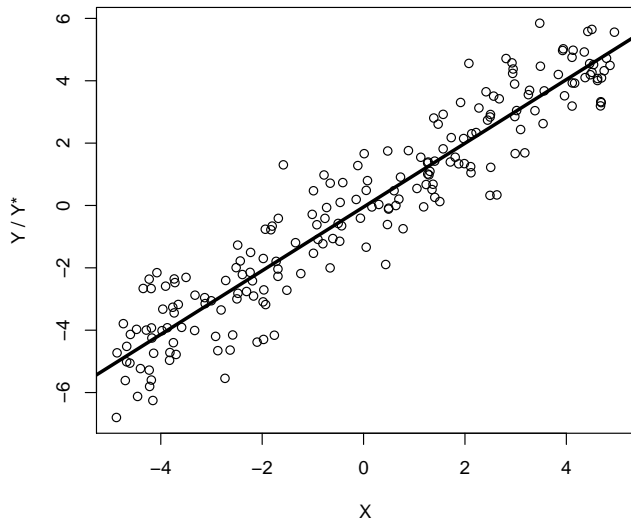
$$\begin{aligned} Y_i &= Y_i^* \text{ if } Y_i^* > L \\ &= L \text{ if } Y_i^* \leq L \end{aligned}$$

“Upper-censored”:

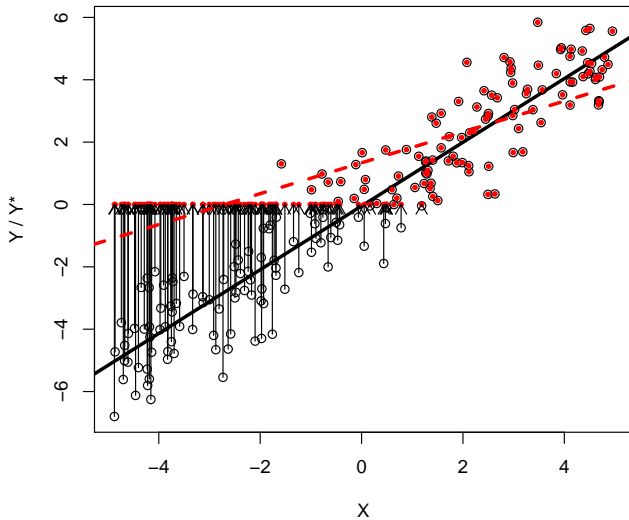
$$\begin{aligned} Y_i &= Y_i^* \text{ if } Y_i^* < U \\ &= U \text{ if } Y_i^* \geq U \end{aligned}$$

→ bias in $\hat{\beta}$ (toward zero) + inconsistency...

Censoring Bias



Censoring Bias



In the lower-censoring case, for $Y^* > L$, we have:

$$\mathbf{L}_1(\beta, \sigma^2 | Y, L) = \prod_{Y_i > L} \phi(Y_i^* | \mathbf{X}_i, \beta, \sigma^2).$$

and for $Y^* \leq L$:

$$\begin{aligned} \Pr(Y_i = L) &= \Pr(Y_i^* \leq L) \\ &= \int_{-\infty}^L \phi(Y_i^* | \mathbf{X}_i, \beta, \sigma^2) dY^* \\ &= \Phi(L | \mathbf{X}_i, \beta, \sigma^2). \end{aligned}$$

which implies:

$$\mathbf{L}_2(\beta, \sigma^2 | Y, L) = \prod_{Y_i = L} \Phi(L | \mathbf{X}_i, \beta, \sigma^2).$$

Combined likelihood:

$$\mathbf{L}(\beta, \sigma^2 | Y, L) = \prod_{Y_i > L} \phi(Y_i^* | \mathbf{X}_i, \beta, \sigma^2) \prod_{Y_i = L} \Phi(L | \mathbf{X}_i, \beta, \sigma^2).$$

One-way unit effects:

$$Y_i^* = \mathbf{X}_i\boldsymbol{\beta} + \alpha_i + u_i$$

Models:

- No fixed-effects conditioning (a la logit) \rightarrow inconsistency.
- Generally use random effects (via `survival` or `xttobit`).

Event Counts: Unit Effects

$$Y_{it} \sim \text{Poisson}(\mu_{it} = \alpha_i \lambda_{it})$$

with $\lambda_{it} = \exp(\mathbf{X}_{it}\beta)$ implies:

$$\begin{aligned} E(Y_{it} \mid \mathbf{X}_{it}, \alpha_i) &= \mu_{it} \\ &= \alpha_i \exp(\mathbf{X}_{it}\beta) \\ &= \exp(\delta_i + \mathbf{X}_{it}\beta) \end{aligned}$$

where $\delta_i = \ln(\alpha_i)$.

- No “incidental parameters” problem (see e.g. Cameron and Trivedi, pp. 281-2)
- Means “brute force” approach works
- Via `xtpoisson` (and `xtnbreg`) in Stata, `glmmML` in R

$$\begin{aligned}\Pr(Y_{i1} = y_{i1}, \dots, Y_{iT} = y_{iT}) &= \int_0^\infty \Pr(Y_{i1} = y_{i1}, \dots, Y_{iT} = y_{iT}) f(\alpha_i) d\alpha_i \\ &= \int_0^\infty \left[\prod_{t=1}^T \Pr(Y_{it} | \alpha_i) \right] f(\alpha_i) d\alpha_i\end{aligned}$$

- Simplest to assume $\alpha_i \sim \Gamma(\theta)$
- Yields a model with $E(Y_{it}) = \lambda_{it}$ and $\text{Var}(Y_{it}) = \lambda_{it} + \frac{\lambda_{it}^2}{\theta}$
- Via `xtpois`, `re` in Stata and `glmmML` or `glmer` in R
- \exists random effects negative binomial too...

R:

- Tobit = `censReg` (in **`censReg`**)
- Poisson (random effects) = `glmmML` in **`glmmML`** or `glmer` in **`lme4`**
- Poisson (fixed effects) = `glmmML` or “brute force”

Stata:

- Tobit = `xttobit` (re only)
- Poisson / negative binomial = `xtpoisson`, `xtnbreg` (both with `fe`, `re` options)
- See notes for more details / examples

Example: State Failure Task Force

```
> summary(SFTF)
```

countryid	year	sftprev	sftpeth	sftpreg
AFG : 9	Min. :1957	Min. :0.0	Min. :0.00	Min. :0.00
ALB : 9	1st Qu.:1967	1st Qu.:0.0	1st Qu.:0.00	1st Qu.:0.00
ARG : 9	Median :1977	Median :0.0	Median :0.00	Median :0.00
AUL : 9	Mean :1979	Mean :0.1	Mean :0.13	Mean :0.12
AUS : 9	3rd Qu.:1992	3rd Qu.:0.0	3rd Qu.:0.00	3rd Qu.:0.00
BEL : 9	Max. :1997	Max. :1.0	Max. :1.00	Max. :1.00

```
(Other):1149
```

sftpgen	poldurab	unuurbpc	ciob	cioc
Min. :0.00	Min. : 0	Min. : 2	Min. : 0	Min. : 0.0
1st Qu.:0.00	1st Qu.: 4	1st Qu.: 23	1st Qu.:14	1st Qu.: 2.0
Median :0.00	Median :12	Median : 41	Median :19	Median : 5.0
Mean :0.08	Mean :21	Mean : 43	Mean :19	Mean : 5.6
3rd Qu.:0.00	3rd Qu.:30	3rd Qu.: 62	3rd Qu.:24	3rd Qu.: 8.0
Max. :1.00	Max. :97	Max. :100	Max. :38	Max. :24.0
	NA's :5	NA's :57		

POLITY	SumEvents
Min. : -10.0	Min. : 0
1st Qu.: -7.0	1st Qu.: 0
Median : -4.0	Median : 0
Mean : -0.7	Mean : 6
3rd Qu.: 8.0	3rd Qu.: 5
Max. : 10.0	Max. :61
NA's :14	NA's :9

```
> pdim(SFTF)
```

```
Unbalanced Panel: n=170, T=1-9, N=1203
```

Panel Tobit: R (see [here](#))

```
> library(plm)
> SFTF.panel<-pdata.frame(SFTF,i="countryid")
> library(censReg)
> Tobit.panel<-censReg(SumEvents~POLITY+unuurbpc+poldurab+year,
+                       data=SFTF.panel,method="BHHH")
> summary(Tobit.panel)
```

Call:

```
censReg(formula = SumEvents ~ POLITY + unuurbpc + poldurab +
        year, data = SFTF.panel, method = "BHHH")
```

Observations:

Total	Left-censored	Uncensored	Right-censored
1132	707	425	0

Coefficients:

	Estimate	Std. error	t value	Pr(> t)
(Intercept)	-1385.21151	60.10481	-23.047	< 2e-16 ***
POLITY	-0.58977	0.09008	-6.547	5.87e-11 ***
unuurbpc	-0.31374	0.03263	-9.616	< 2e-16 ***
poldurab	-0.34628	0.02624	-13.198	< 2e-16 ***
year	0.70470	0.03048	23.121	< 2e-16 ***
logSigmaMu	2.83694	0.05035	56.341	< 2e-16 ***
logSigmaNu	2.58187	0.02160	119.522	< 2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

BHHH maximisation, 40 iterations

Return code 2: successive function values within tolerance limit

Log-likelihood: -2020 on 7 Df

Panel Poisson (Random Effects)

```
> library(lme4)
> Poisson.RE<-glmer(ciob~POLITY+unuurbpc+poldurab+I(year-1900)+(1|countryid),
  data=SFTF,family="poisson")
> summary(Poisson.RE)
Generalized linear mixed model fit by maximum likelihood (Laplace
Approximation) [glmerMod]
Family: poisson ( log )
Formula:
ciob ~ POLITY + unuurbpc + poldurab + I(year - 1900) + (1 | countryid)
Data: SFTF

      AIC      BIC   logLik deviance df.resid
  6811    6841    -3399     6799     1126

Random effects:
Groups   Name      Variance Std.Dev.
countryid (Intercept) 0.159   0.399
Number of obs: 1132, groups:  countryid, 160

Fixed effects:
              Estimate Std. Error z value    Pr(>|z|)
(Intercept)   1.200274   0.063085   19.03    < 2e-16 ***
POLITY        -0.003484   0.001812   -1.92     0.055 .
unuurbpc       0.005996   0.001064    5.64 0.000000017 ***
poldurab       0.001167   0.000672    1.74     0.082 .
I(year - 1900) 0.016385   0.000855   19.16    < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
              (Intr) POLITY unurbpc poldrb
POLITY         0.354
unuurbpc      -0.075 -0.139
poldurab       0.224  0.348 -0.087
I(yer-1900)   -0.628 -0.273 -0.589 -0.313
convergence code: 0
Model failed to converge with max|grad| = 0.00158426 (tol = 0.001, component 1)
Model is nearly unidentifiable: very large eigenvalue
- Rescale variables?
```

Panel Poisson (Random Effects – Alternative)

```
> library(glmML)
> Poisson.RE.alt<-glmML(cio~POLITY+unuurbpc+poldurab+I(year-1900),
                        data=SFTF,cluster=countryid,
                        family="poisson")
> summary(Poisson.RE.alt)
```

```
Call: glmML(formula = cio ~ POLITY + unuurbpc + poldurab + I(year - 1900),
             family = "poisson", data = SFTF, cluster = countryid)
```

	coef	se(coef)	z	Pr(> z)
(Intercept)	1.20027	0.063120	19.02	0.000000000
POLITY	-0.00348	0.001814	-1.92	0.055000000
unuurbpc	0.00600	0.001064	5.63	0.000000018
poldurab	0.00117	0.000672	1.74	0.082000000
I(year - 1900)	0.01639	0.000856	19.15	0.000000000

```
Scale parameter in mixing distribution: 0.399 gaussian
Std. Error: 0.0263
```

```
LR p-value for H_0: sigma = 0: 2.28e-289
```

```
Residual deviance: 1590 on 1126 degrees of freedom AIC: 1600
```

Panel Poisson (Fixed Effects – “brute force”)

```
> Poisson.FE<-glm(cio~POLITY+unuurbpc+poldurab+I(year-1900)+
  as.factor(countryid),data=SFTF,family="poisson")
> summary(Poisson.FE)
```

Call:

```
glm(formula = cio ~ POLITY + unuurbpc + poldurab + I(year -
  1900) + as.factor(countryid), family = "poisson", data = SFTF)
```

Deviance Residuals:

	Min	1Q	Median	3Q	Max
	-4.806	-0.312	0.069	0.364	2.863

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	1.040769	0.117296	8.87	< 2e-16 ***
POLITY	-0.007437	0.001939	-3.84	0.00013 ***
unuurbpc	0.005011	0.001580	3.17	0.00151 **
poldurab	-0.000477	0.000749	-0.64	0.52386
I(year - 1900)	0.018411	0.001115	16.51	< 2e-16 ***
as.factor(countryid)ALB	-0.376632	0.142587	-2.64	0.00826 **
as.factor(countryid)ALG	0.200591	0.131453	1.53	0.12702
.				
.				
.				
as.factor(countryid)ZAM	0.094994	0.132209	0.72	0.47244
as.factor(countryid)ZIM	-0.053680	0.137511	-0.39	0.69627

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1

(Dispersion parameter for poisson family taken to be 1)

Null deviance: 4453.85 on 1131 degrees of freedom
Residual deviance: 942.45 on 968 degrees of freedom
(71 observations deleted due to missingness)
AIC: 6483

Number of Fisher Scoring iterations: 5