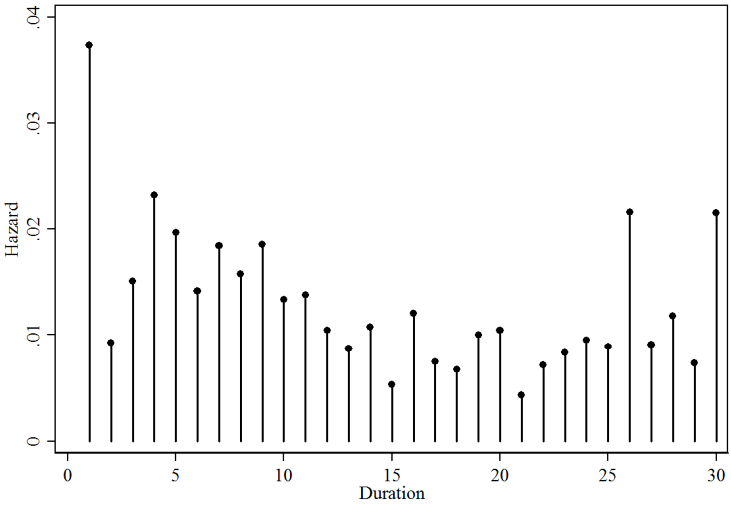
ICPSR 2017 "Advanced Maximum Likelihood": Survival Analysis

Day Four

August 10, 2017

The Discrete-Time Idea



A General Discrete-Time Model

Process:

$$t \in \{1, 2, ...t_{\mathsf{max}}\}$$

Density:

$$f(t) = \Pr(T = t)$$

CDF:

$$\begin{aligned} F(t) &= \Pr(T \leq t) \\ &= \sum_{j=1}^{t} f(t_j) \end{aligned}$$

A General Discrete-Time Model

Survival function:

$$
\begin{aligned}
S(t) &\equiv \Pr(T \geq t) \\
&= 1 - F(t) \\
&= \sum_{j=t}^{t_{\text{max}}} f(t_j)
\end{aligned}
$$

Hazard function:

$$
\begin{aligned}
h(t) &\equiv \Pr(T = t | T \geq t) \\
&= \frac{f(t)}{S(t)}
\end{aligned}
$$

A General Discrete-Time Model

Conditional Pr(Survival):

$$\Pr(T > t | T \geq t) = 1 - h(t)$$

Implies:

$$\begin{aligned}
S(t) &= \Pr(T > t | T \geq t) \times \Pr(T > t-1 | T \geq t-1) \times \Pr(T > t-2 | T \geq t-2) \times ... \\
&\quad \times \Pr(T > 2 | T \geq 2) \times \Pr(T > 1 | T \geq 1) \\
&= [1 - h(t)] \times [1 - h(t-1)] \times [1 - h(t-2)] \times ... \times [1 - h(2)] \times [1 - h(1)] \\
&= \prod_{j=0}^{t} [1 - h(t-j)]
\end{aligned}$$

A General Discrete-Time Model

which means:

$$
\begin{aligned}
f(t) &= h(t)S(t) \\
&= h(t) \times [1 - h(t-1)] \times [1 - h(t-2)] \times \ldots \\
&\quad \times [1 - h(2)] \times [1 - h(1)] \\
&= h(t) \prod_{j=1}^{t-1} [1 - h(t-j)]
\end{aligned}
$$

General Discrete-Time Model: Likelihood

$$L = \prod_{i=1}^{N} \left\{ h(t) \prod_{j=1}^{t-1} [1 - h(t-j)] \right\}^{Y_{it}} \left\{ \prod_{j=0}^{t} [1 - h(t-j)] \right\}^{1-Y_{it}}$$

Ordered-Categorical Models

For K small:

$$\Pr(T_i \leq k) = \frac{\exp(\tau_k - \mathbf{X}_i\beta)}{1 + \exp(\tau_k - \mathbf{X}_i\beta)}$$

$$\ln\left[\frac{\Pr(T_i \leq \kappa)}{\Pr(T_i > \kappa)}\right] = \tau_\kappa - \mathbf{X}_i\beta$$

Grouped-Data ("BTSCS") Approaches

$$\Pr(Y_{it} = 1) = f(\mathbf{X}_{it}\beta)$$

- logit
- probit
- c-log-log
- etc.

BTSCS: Advantages

- Easily estimated, interpreted and understood
- Natural interpretations:
  - $\hat{\beta}_0 \approx$ "baseline hazard"
  - Covariates shift this up or down.
- Can incorporate data in time-varying covariates
- Lots of software

(Potential) Disadvantages

- Requires time-varying data
- *Must deal with time dependence explicitly*

Temporal Issues in Grouped-Data Models

(Implicit) "Baseline" hazard:

$$h_0(t) = \frac{\exp(\beta_0)}{1 + \exp(\beta_0)}$$

$\longrightarrow$ No temporal dependence / "flat" hazard

Temporal Issues in Grouped-Data Models

Time trend:

$$\Pr(Y_{it} = 1) = f(\mathbf{X}_{it}\beta + \gamma T_{it})$$

- $\hat{\gamma} > 0 \ \rightarrow$ rising hazard
- $\hat{\gamma} < 0 \ \rightarrow$ declining hazard
- $\hat{\gamma} = 0 \ \rightarrow$ "flat" (exponential) hazard

Variants/extensions: Polynomials...

$$\Pr(Y_{it} = 1) = f(\mathbf{X}_{it}\beta + \gamma_1 T_{it} + \gamma_2 T_{it}^2 + \gamma_3 T_{it}^3 + ...)$$

Temporal Issues in Grouped-Data Models

"Time dummies":

$$\Pr(Y_{it} = 1) = f[\mathbf{X}_{it}\beta + \alpha_1 I(T_{i1}) + \alpha_2 I(T_{i2}) + ... + \alpha_{t_{\max}} I(T_{it_{\max}})]$$

$\longrightarrow$ BKT's cubic splines; might also use:

- Fractional polynomials
- Smoothed duration
- Loess/lowess fits
- Other splines (B-splines, P-splines, natural splines, etc.)

Discrete-Time Model Selection

- Theory
- Formal tests
- Fitted values

Equivalency One: Cox $\equiv$ Conditional Logit

$$\Pr(Y_i = j) = \frac{\exp(\mathbf{X}_{ij}\beta + \mathbf{Z}_j\gamma)}{\sum_{\ell=1}^{J} \exp(\mathbf{X}_{i\ell}\beta + \mathbf{Z}_\ell\gamma)}$$

$$\Pr(Y_i = j) = \frac{\exp(\mathbf{X}_{ij}\beta)}{\sum_{\ell=1}^{J} \exp(\mathbf{X}_{i\ell}\beta)}$$

$$L_k = \frac{\exp(\mathbf{X}_k\beta)}{\sum_{\ell \in R_j} \exp(\mathbf{X}_\ell\beta)}.$$

The point: Cox $\equiv$ Conditional logit

Cox-Poisson Equivalence

Grouped-data duration models and the continuous-time Cox model are equivalent.

Cox-Poisson Equivalence

Cox:

$$S_i(t) = \exp\left[-\exp(\mathbf{X}_i\beta)\int_0^t h_0(t)\,dt\right]$$

Poisson:

$$\Pr(Y = y) = \frac{\exp(-\lambda)\lambda^y}{y!}$$

$$\begin{aligned}\Pr(Y_{it} = 0) &= \exp(-\lambda) \\ &= \exp[-\exp(\mathbf{X}_i\beta)]\end{aligned}$$

Example: Oneal & Russett (1950-1985)

No time variable / “flat” hazard:

```
> OR.logit<-glm(dispute~allies+contig+capratio+growth+democracy+trade,
              data=OR,na.action=na.exclude,family="binomial")
> summary(OR.logit)

Coefficients:
             Estimate Std. Error z value Pr(>|z|)
(Intercept)  -4.32668    0.11451 -37.785  < 2e-16 ***
allies       -0.47969    0.11275  -4.255 2.09e-05 ***
contig        1.35358    0.12091  11.195  < 2e-16 ***
capratio     -0.19620    0.05011  -3.916 9.01e-05 ***
growth       -3.42753    1.25181  -2.738  0.00618 **
democracy    -0.40120    0.10063  -3.987 6.70e-05 ***
trade       -21.07611   11.30396  -1.864  0.06225 .
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
```

Example, Continued

Linear trend:

```
> OR$duration<-OR$stop
> OR.trend<-glm(dispute~allies+contig+capratio+growth+democracy+trade
               +duration,data=OR,na.action=na.exclude,family="binomial")
> summary(OR.trend)

Coefficients:
             Estimate Std. Error z value Pr(>|z|)
(Intercept) -3.271136   0.134709 -24.283  < 2e-16 ***
allies      -0.362966   0.114140  -3.180 0.001473 **
contig       0.996908   0.123978   8.041 8.91e-16 ***
capratio    -0.235655   0.052763  -4.466 7.96e-06 ***
growth      -3.957428   1.225716  -3.229 0.001244 **
democracy   -0.361150   0.099515  -3.629 0.000284 ***
trade       -2.870981   9.861298  -0.291 0.770947
duration    -0.091189   0.008098 -11.260  < 2e-16 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
```

Example, Continued

Fourth-Order polynomial trend:

```
OR$d2<-OR$duration^2*0.1
OR$d3<-OR$duration^3*0.01
OR$d4<-OR$duration^4*0.001

OR.P4<-glm(dispute~allies+contig+capratio+growth+democracy+trade
              +duration+d2+d3+d4,data=OR,na.action=na.exclude,
              family="binomial")

> summary(OR.P4)

Coefficients:
             Estimate Std. Error z value Pr(>|z|)
(Intercept) -3.401363   0.206815 -16.446  < 2e-16 ***
allies      -0.364127   0.114201  -3.188  0.00143 **
contig       0.995584   0.124074   8.024 1.02e-15 ***
capratio    -0.228355   0.052257  -4.370 1.24e-05 ***
growth      -3.864329   1.245617  -3.102  0.00192 **
democracy   -0.392457   0.100693  -3.898 9.72e-05 ***
trade       -4.032292   9.631171  -0.419  0.67546
duration     0.058036   0.091465   0.635  0.52574
d2          -0.274958   0.128454  -2.141  0.03231 *
d3           0.136086   0.063230   2.152  0.03138 *
d4          -0.018863   0.009914  -1.903  0.05709 .
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
```

Polynomial Improvement?

```
> P4test
Analysis of Deviance Table

Model 1: dispute ~ allies + contig + capratio + growth + democracy + trade
Model 2: dispute ~ allies + contig + capratio + growth + democracy + trade +
    duration + d2 + d3 + d4

  Resid. Df Resid. Dev Df Deviance  Pr(>Chi)
1     20441     3693.8
2     20437     3510.0  4   183.76 < 2.2e-16 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
```

Example: "Time Dummies"

"Time dummies":

```
> OR.dummy<-glm(dispute~allies+contig+capratio+growth+democracy+trade
           +as.factor(duration),data=OR,na.action=na.exclude,
           family="binomial")

> summary(OR.dummy)

Coefficients:
                       Estimate Std. Error z value Pr(>|z|)
(Intercept)            -3.61115    0.18219 -19.820  < 2e-16 ***
allies                 -0.36922    0.11441  -3.227 0.001251 **
contig                  0.99389    0.12417   8.005 1.20e-15 ***
capratio               -0.22778    0.05219  -4.364 1.27e-05 ***
growth                 -3.97619    1.24940  -3.182 0.001460 **
democracy              -0.39559    0.10077  -3.926 8.65e-05 ***
trade                  -3.46727    9.62606  -0.360 0.718700
as.factor(duration)2    0.45489    0.19606   2.320 0.020331 *
as.factor(duration)3    0.36020    0.20632   1.746 0.080843 .
as.factor(duration)4    0.14188    0.22175   0.640 0.522289

   <output omitted>

as.factor(duration)33  -1.64467    1.01715  -1.617 0.105891
as.factor(duration)34  -0.86966    0.73158  -1.189 0.234541
as.factor(duration)35  -1.38777    1.01857  -1.362 0.173049
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
```

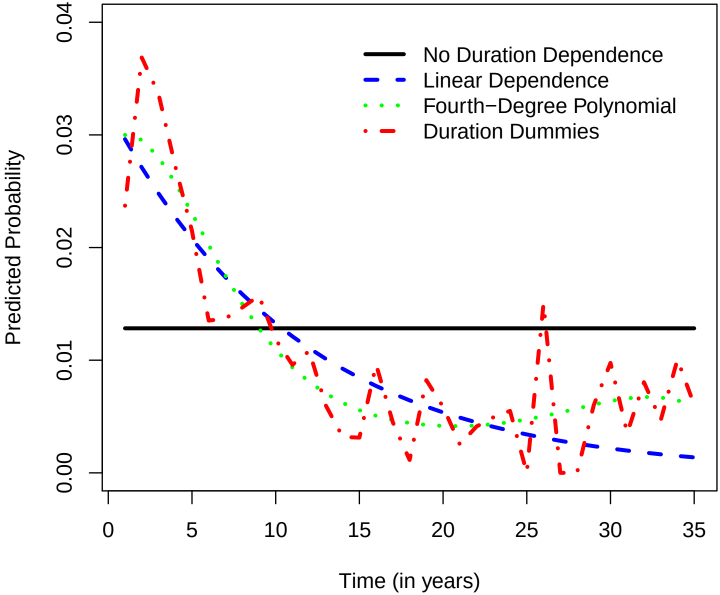
“Time Dummies,” continued

```
> Test.Dummies<-anova(OR.logit,OR.dummy,test="Chisq")

> Test.Dummies
Analysis of Deviance Table

Model 1: dispute ~ allies + contig + capratio + growth + democracy + trade
Model 2: dispute ~ allies + contig + capratio + growth + democracy + trade +
    as.factor(duration)
  Resid. Df Resid. Dev Df Deviance  Pr(>Chi)
1     20441     3693.8
2     20407     3464.4 34   229.38 < 2.2e-16 ***
```

Predicted “Hazards” (Probabilities)



Cox / Poisson Equivalence

Cox model:

```
OR.Cox<-coxph(Surv(OR$start,OR$stop,OR$dispute)~allies+contig+capratio+
              growth+democracy+trade,data=OR,method="breslow")
```

Poisson:

```
OR.Poisson<-glm(dispute~allies+contig+capratio+growth+democracy+trade
              +as.factor(duration),data=OR,na.action=na.exclude,
              family="poisson")
```