GSERM - Oslo 2019
Cox and Discrete-Time Models

January 10, 2019 (afternoon session)

Cox (1972)

h_0 := baseline hazard.

Basic idea:

$$h_i(t) = h_0(t)\exp(\mathbf{X}_i\beta)$$

Note:

- $h_0(t) \equiv h(t|\mathbf{X} = 0)$
- Changes in $\mathbf{X}$ shift $h(t)$ *proportionally*

model of proportional hazards.

$$\begin{aligned}
\text{HR} &= \frac{h_0(t)\exp(X_1\hat{\beta})}{h_0(t)\exp(X_0\hat{\beta})} \\
&= \exp[(1-0)\hat{\beta}] \\
&= \exp(\hat{\beta})
\end{aligned}$$

Also, because

$$S(t) = \exp[-H(t)]$$

then

$$\begin{aligned}
S(t) &= \exp\left[-\int_0^t h(t)\,dt\right] \\
&= \exp\left[-\exp(\mathbf{X}_i\beta)\int_0^t h_0(t)\,dt\right] \\
&= \left[\exp\left(-\int_0^t h_0(t)\,dt\right)\right]^{\exp(\mathbf{X}_i\beta)} \\
&= [S_0(t)]^{\exp(\mathbf{X}_i\beta)}
\end{aligned}$$

goal: do not make distributional assumptions of h_0

Assume N_C distinct event times t_j, with no "ties."

Then:

$$
\begin{aligned}
&\Pr(\text{Individual } k \text{ experienced the event at } t_j \mid \textit{One} \text{ observation experienced the event at } t_j) \\
&\quad = \frac{\Pr(\text{At-risk observation } k \text{ experiences the event of interest at } t_j)}{\Pr(\text{One at-risk observation experiences the event of interest at } t_j)} \\
&\quad = \frac{h_k(t_j)}{\sum_{\ell \in R_j} h_\ell(t_j)}
\end{aligned}
$$

R_j all individuals at risk time t_j

and hazard rate of h_k for individual k

proportional hazrad risk for individual.

the ratio is actually a probability for individual k to get hazard exactly at time t_j

(similar like likelihood)

Partial Likelihood (continued)

$$
\begin{aligned}
L_i &= \frac{h_0(t_j)\exp(\mathbf{X}_i\beta)}{\sum_{\ell \in R_j} h_0(t_j)\exp(\mathbf{X}_\ell\beta)} \\
&= \frac{h_0(t_j)\exp(\mathbf{X}_i\beta)}{h_0(t_j)\sum_{\ell \in R_j} \exp(\mathbf{X}_\ell\beta)} \\
&= \frac{\exp(\mathbf{X}_i\beta)}{\sum_{\ell \in R_j} \exp(\mathbf{X}_\ell\beta)}
\end{aligned}
$$

partial likelihood, coz no distributional assumptions.

assumptions are just, everyone shares a baseline hazard rate and covariates shift this

$$
L = \prod_{i=1}^{N} \left[\frac{\exp(\mathbf{X}_i\beta)}{\sum_{\ell \in R_j} \exp(\mathbf{X}_\ell\beta)} \right]^{C_i}
$$

$$
\ln L = \sum_{i=1}^{N} C_i \left\{ \mathbf{X}_i\beta - \ln\left[\sum_{\ell \in R_j} \exp(\mathbf{X}_\ell\beta) \right] \right\}
$$

Notes on Partial Likelihood

- PL is

consistent estimator for beta

  - Consistent
  - Asymptotically normal
  - Slightly inefficient (but asymptotically efficient)

- Considers order of events, but not actual duration
- Censored events: Modify R_j
- No ties

see comments on paper: the order of events is important, not the duration; asymptotically equal.

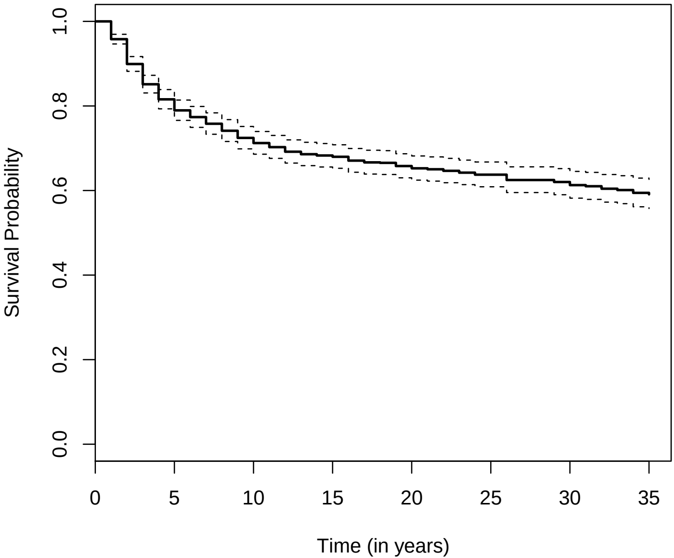
Example: Interstate War, 1950-1985

- Dyad-years for "politically-relevant" dyads
- $N = 827$, $NT = 20448$.
- Covariates:
  - Whether $(=1)$ or not the two countries are *allies*,
  - Whether $(=1)$ or not the two countries are *contiguous*,
  - The *capability ratio* of the two countries,
  - The lower of the two countries' (GDP) *growth* (rescaled),
  - The lower of the two countries' *democracy* (POLITY IV) scores (rescaled to [-1,1]), and
  - The amount of *trade* between the two countries, as a fraction of joint GDP.

```
> summary(OR)
     dyadid             year          start            stop           futime     
 Min.   :  2020   Min.   :1951   Min.   : 0.00   Min.   : 1.00   Min.   : 5.00  
 1st Qu.:100365   1st Qu.:1965   1st Qu.: 5.00   1st Qu.: 6.00   1st Qu.:23.00  
 Median :220235   Median :1972   Median :11.00   Median :12.00   Median :31.00  
 Mean   :253305   Mean   :1971   Mean   :12.32   Mean   :13.32   Mean   :28.97  
 3rd Qu.:365600   3rd Qu.:1979   3rd Qu.:19.00   3rd Qu.:20.00   3rd Qu.:35.00  
 Max.   :900920   Max.   :1985   Max.   :34.00   Max.   :35.00   Max.   :35.00  
    dispute            allies           contig           trade        
 Min.   :0.00000   Min.   :0.0000   Min.   :0.0000   Min.   :0.00000  
 1st Qu.:0.00000   1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:0.00000  
 Median :0.00000   Median :0.0000   Median :0.0000   Median :0.00020  
 Mean   :0.01981   Mean   :0.3563   Mean   :0.3099   Mean   :0.00231  
 3rd Qu.:0.00000   3rd Qu.:1.0000   3rd Qu.:1.0000   3rd Qu.:0.00120  
 Max.   :1.00000   Max.   :1.0000   Max.   :1.0000   Max.   :0.17680  
     growth             democracy          capratio      
 Min.   :-0.264900   Min.   :-1.0000   Min.   : 0.0100  
 1st Qu.:-0.004800   1st Qu.:-0.8000   1st Qu.: 0.0462  
 Median : 0.014700   Median :-0.7000   Median : 0.2220  
 Mean   : 0.007823   Mean   :-0.3438   Mean   : 1.6677  
 3rd Qu.: 0.027800   3rd Qu.: 0.2000   3rd Qu.: 1.1560  
 Max.   : 0.164700   Max.   : 1.0000   Max.   :78.9296  
```

dyads with overlapping memberships can cause issues (who would go into war with multiple countries) := dyadic dependence. However, there are explicit means (network effect models) to deal with this.

The Data (Kaplan-Meier plot)



lots of censored cases here, i.e. many countries do not go to war.

Software

R:

- `coxph` in `survival` (preferred)
- `cph` in `design`
- Plots: `plot(survfit(PHobject))`

Stata:

- Basic command = `stcox`
- `stset` first
- Options: `robust`, various methods for ties, postestimation commands

Model Fitting

```
> ORCox.br<-coxph(OR.S~allies+contig+capratio+growth+democracy+trade,
                  data=OR,na.action=na.exclude,method="breslow")

> summary(ORCox.br)

  n= 20448, number of events= 405

              coef exp(coef) se(coef)      z Pr(>|z|)
allies    -0.34849   0.70576  0.11096 -3.141 0.001686 **
contig     0.94861   2.58213  0.12173  7.793 6.55e-15 ***
capratio  -0.22303   0.80009  0.05164 -4.319 1.57e-05 ***
growth    -3.69487   0.02485  1.19950 -3.080 0.002068 **
democracy -0.38194   0.68254  0.09915 -3.852 0.000117 ***
trade     -3.22857   0.03961  9.45588 -0.341 0.732776
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
.
.
.
```

Model Fitting (continued)

```
.
.
.
          exp(coef) exp(-coef) lower .95 upper .95
allies      0.70576     1.4169 5.678e-01 8.772e-01
contig      2.58213     0.3873 2.034e+00 3.278e+00
capratio    0.80009     1.2499 7.231e-01 8.853e-01
growth      0.02485    40.2402 2.368e-03 2.608e-01
democracy   0.68254     1.4651 5.620e-01 8.289e-01
trade       0.03961    25.2436 3.540e-10 4.433e+06

Concordance= 0.714  (se = 0.015 )
Rsquare= 0.01   (max possible= 0.234 )
Likelihood ratio test= 210.3  on 6 df,   p=0
Wald test            = 159.8  on 6 df,   p=0
Score (logrank) test = 185.8  on 6 df,   p=0
```

Rsquared is misleading and probably should not be reported.

Interpretation: Hazard Ratios

$$HR = \exp[(\mathbf{X}_j - \mathbf{X}_k)\hat{\beta}]$$

Means:

- $HR = 1 \;\leftrightarrow\; \hat{\beta} = 0$
- $HR > 1 \;\leftrightarrow\; \hat{\beta} > 0$
- $HR < 1 \;\leftrightarrow\; \hat{\beta} < 0$

$$\text{Percentage difference} = 100 \times \{\exp[(\mathbf{X}_j - \mathbf{X}_k)\hat{\beta}] - 1\}.$$

Example: Hazard Ratios

From above:

	exp(coef)	exp(-coef)	lower .95	upper .95
allies	0.70576	1.4169	5.678e-01	8.772e-01
contig	2.58213	0.3873	2.034e+00	3.278e+00
capratio	0.80009	1.2499	7.231e-01	8.853e-01
growth	0.02485	40.2402	2.368e-03	2.608e-01
democracy	0.68254	1.4651	5.620e-01	8.289e-01
trade	0.03961	25.2436	3.540e-10	4.433e+06

Interpretation:

- Countries which are *allies* have an expected $(0.706 - 1) \times 100) = 29.4$ percent lower hazard of conflict than those that are not.
- *Contiguous* countries have $(2.582 - 1) \times 100 = 158$ percent higher hazards of conflict than non-contiguous ones.
- A one-unit increase in *democracy* corresponds to a $(0.683 - 1) \times 100 = 31.7$ percent decrease in the expected hazard of conflict.

one unit change in economic growth leads to around -98% change in expected hazard of conflict.

but growth in one unit change is not a reasonable scale.

Hazard Ratios: Scaling Covariates

It is good for one-unit changes to be meaningful / realistic...

```
> OR$growthPct<-OR$growth*100
> summary(coxph(OR.S~allies+contig+capratio+growthPct+democracy+trade,
                data=OR,na.action=na.exclude, method="breslow"))
.
.
.
          exp(coef) exp(-coef) lower .95 upper .95
allies      0.70576     1.4169 5.678e-01 8.772e-01
contig      2.58213     0.3873 2.034e+00 3.278e+00
capratio    0.80009     1.2499 7.231e-01 8.853e-01
growthPct   0.96373     1.0376 9.413e-01 9.867e-01
democracy   0.68254     1.4651 5.620e-01 8.289e-01
trade       0.03961    25.2436 3.540e-10 4.433e+06
```

Note:

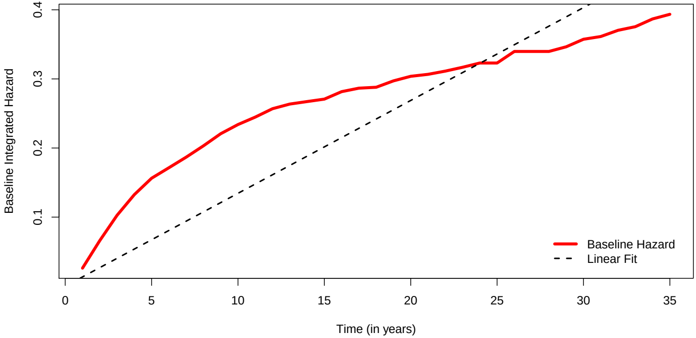
- Previous HR for `growth` $= 0.02485 \rightarrow 97.5$ percent decrease in $\hat{h}(t)$
- HR for `growthPct` is now 0.964; 1 unit increase $\rightarrow$ 4% decrease in $\hat{h}(t)$
- Same result, proportionally: $0.96373^{100} = 0.02485$

Baseline Hazards

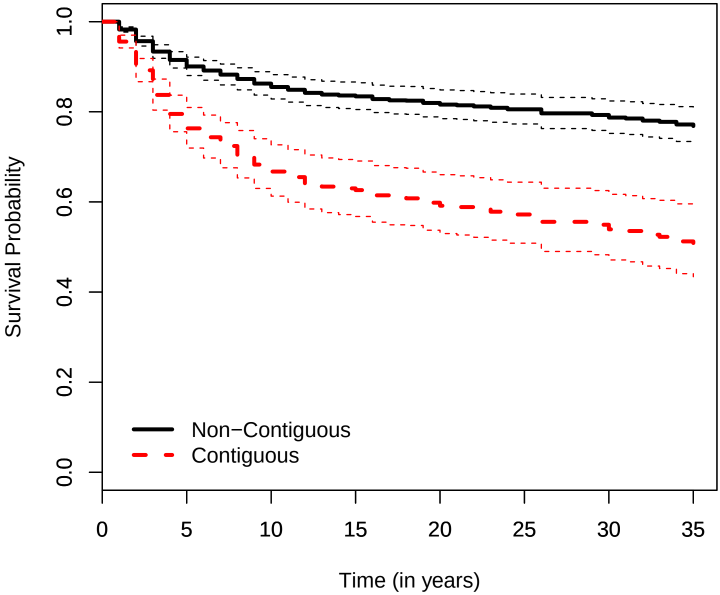
Because the Cox model is semiparametric, it uses a conventional / univariate (Nelson-Aalen) estimate of the "baseline" hazard:

```
OR.BH<-basehaz(ORCox.br,centered=FALSE)
```

empirical baseline integrated hazard.



Comparing Survival Curves



$\exists$ ties...

Their presence biases Cox $\hat{\beta}$s toward zero.

- Call $d_j > 0$ the number of events occurring at t_j, and
- D_j the set of d_j observations that have the event at t_j.

more ties means less unique durations with events, leading to less variation,
in extreme case no more variation, thus beta zero.

if d_j, same as cox model; if d_j > 1, adjust for events

"good enough" approximation,

but Efron figured it is systematically wrong.

Means of handling ties:

- Breslow:

$$L_{\text{Breslow}}(\beta) = \prod_{i=1}^{N} \frac{\exp\left[\left(\sum_{q \in D_j} \mathbf{X}_q\right)\beta\right]}{\left[\sum_{\ell \in R_j} \exp(\mathbf{X}_\ell \beta)\right]^{d_j}}$$

- Efron

("bootstrapping")

$$\ln L_{\text{Efron}}(\beta) = \sum_{j=1}^{J} \sum_{i \in D_j} \left\{ \mathbf{X}_i \beta - \frac{1}{d_j} \sum_{k=1}^{d_j - 1} \ln \left[\sum_{\ell \in R_j} \exp(\mathbf{X}_\ell \beta) - k \left(\frac{1}{d_j} \sum_{\ell \in D_j} \exp(\mathbf{X}_\ell \beta) \right) \right] \right\}$$

add a little bit of noise to each of the durations, so we have an untied order, then simulate many times and take average.

Ties (continued)

- "Exact" (partial likelihood)

$$\ln L_{\text{Exact}}(\boldsymbol{\beta}) = \sum_{j=1}^{J} \left\{ \sum_{i \in R_j} \delta_{ij}(\mathbf{X}_i \boldsymbol{\beta}) - \ln[f(r_j, d_j)] \right\}$$

where

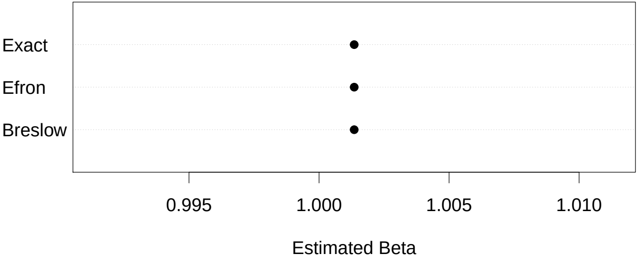
$$\begin{aligned} f(r,d) &= g(r-1,d) + g(r-1,d-1)\exp(\mathbf{X}_k \boldsymbol{\beta}), \\ k &= r\text{th observation in } R_j, \\ r_j &= \text{cardinality of } R_j, \text{ and} \\ g(r,d) &= \begin{cases} 0 \text{ if } r < d, \\ 1 \text{ if } d = 0 \end{cases} \end{aligned}$$

computationally expensive when there are many events at one timepoint.
=> thus the previously proposed approximations.

Ties: Example

```
set.seed(7222009)
Data<-as.data.frame(cbind(c(rep(1,times=400)),
            c(rep(c(0,1),times=200))))
colnames(Data)<-c("C","X")
Data$T<-rexp(400,exp(0+1*Data$X)) # B = 1.0
Data.S<-Surv(Data$T,Data$C)

D.br<-coxph(Data.S~X,data=Data,method="breslow")
D.ef<-coxph(Data.S~X,data=Data,method="efron")
D.ex<-coxph(Data.S~X,data=Data,method="exact")
```

no ties, all methods are the same.



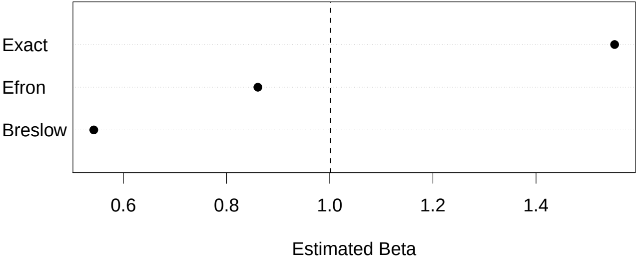
Ties: Example (continued)

```
Data$Tied<-round(Data$T,0)

DataT.S<-Surv(Data$Tied,Data$C)

DT.br<-coxph(DataT.S~X,data=Data,method="breslow")
DT.ef<-coxph(DataT.S~X,data=Data,method="efron")
DT.ex<-coxph(DataT.S~X,data=Data,method="exact")
```

with ties, different results

(rounding compresses variation, coz 0 not possible and everything below is rounded to 1)

so (probably) exact method is closer to the actual beta.



stata defaults to breslow, which is bad !!

Ties: Practical Advice

- All approx. are identical if $\nexists$ ties
- Few ties $=$ similar results
- When ties are present, Breslow $<$ Efron $<$ "Exact" methods
- If you want to learn more about ties in the Cox model, read my paper....

Cox vs. Parametric Models

cox model uses (a little) less information from the observations

Conceptual considerations:

- Theory
- Nature of $h(t)$
- Relative importance: Bias vs. efficiency
- Need / willingness for out-of-sample predictions / forecasting

no theory, no nature of h, bias important, no out of sample predictions => cox model

parametric is good for (out of sample) prediction, cox is only defined for t when an event occured at t

also good when there is substantive theory behind the process and you can make distributional assumptions with confidence.

medicine is interested in no-bias in-sample covariate effect; engineers know the mechanics, they do not care about covariate effect, they want prediction (?)

Cox, On His Model

Reid: "What do you think of the cottage industry that's grown up around [the Cox model]?"

Cox: "In the light of further results one knows since, I think I would normally want to tackle the problem parametrically... I'm not keen on non-parametric formulations normally."

Reid: "So if you had a set of censored survival data today, you might rather fit a parametric model, even though there was a feeling among the medical statisticians that that wasn't quite right."

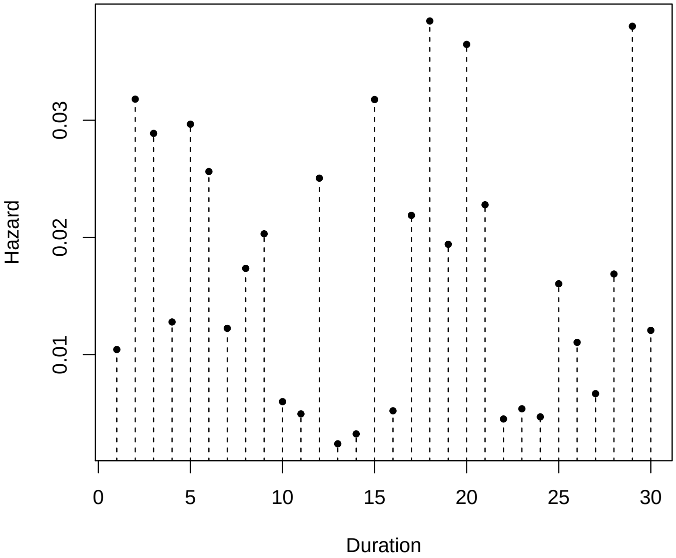
Cox: "That's right, but since then various people have shown that the answers are very insensitive to the parametric formulation of the underlying distribution. And if you want to do things like predict the outcome for a particular patient, it's much more convenient to do that parametrically."

– From Reid (1994).

Extensions

- Cox Models for *repeated events*
- Models with “frailties”
- Competing risks / “cured” subpopulations
- etc.

The Discrete-Time Idea



A General Discrete-Time Model

Process:

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

Density:

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

CDF:

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

Survival function:

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

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:

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

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

Y_it is dummy indicator if the event occured at that timepoint for unit i. Similar to the C_i

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

Ordered-Categorical Models

(viable as alternative for small K,
and it can be interpreted as odds-ratio)

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

just take survival as binary outcome being 1 or 0 using e.g. logistic regression.

- 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:

no t, i.e. the model assumes a flat 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

add T with a coefficient;
looks like a Weibull model.

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...

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

instead of including form of hazard as distributional information, it is explicitly modelled in the equation.

Temporal Issues in Grouped-Data Models

(similar to Cox model)

due to many alpha dummies, overparameterized, less efficient.

"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.)

- Theory
- Formal tests
- Fitted values

Equivalency One: Cox $\equiv$ Conditional Logit

(panel logistic regression for binary outcomes)

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

multinomial logistic regression

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

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

likelihood for cox model

The point: Cox $\equiv$ Conditional logit

the event at a particular time "chooses" one observation according to the observations characteristics from the risk set;
equally in multionomial logistic regression with categorical outcome, where one person chooses according to products characteristics a particular product.

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:

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

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

probability of drawing a zero from a poisson process.

count of how many events happened to that unit at the timepoint
and the zero is the event not happening, i.e. survival.

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
```

agnostic to time,
risk constant at time.

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”

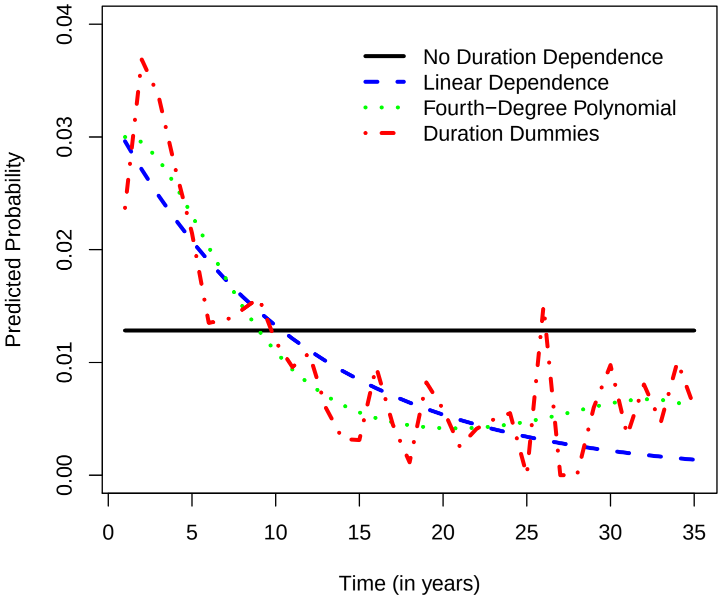
“Time dummies”:

```
> 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")
```

needs breslow

Poisson:

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

duration must be included as dummies to get the variation.