Thrips jumping: using survival analyses and GLM methods for non-medical data

Ruth Butler, Rob van Tol, Melanie Davidson
Thrips

WFT: Western Flower Thrips <3mm
- Important economic pest
- Feeds on: many food and ornamental crops
  -(Capsicum, Chrysanthemum, Bean…….)
- Direct feeding damage; Carries viruses

Objective:
Assess behaviour of female WFT on different host plants with/without added methyl isonicotinate (MI), an allelochemical
Experimental Set-Up

Lab experiment, using 10mm leaf /petal discs

Gas fed from source bottle via tube onto disc

Single thrips placed on disc in a run

Observed for 5min

Recoded:
‘Jumped’ from disc: Yes/ No
Time to ‘Jump’
Time moving before ‘Jumping’
Host-plant experiment

8 treatments:
Gas (MI or air) x
Plant (Bean Leaf, Pepper leaf, Chrysanthemum Leaf, Chrys. petal)

20 replicates done in turn:
each rep. is 8 discs with 1 thrips done sequentially.

systematic treatment order:
Rep 1, 3…:
Bean (C,MI), Pepper (C,MI), ChLeaf (C,MI), ChPetal (C,MI)
reversed for Rep 2, 4…

Potential Random effects: Rep, Order, Individual Thrips
Analysis methods

Percentage jumping, Percentage time moving before jumping:
   Binomial GLM, or, with random effects, GLMM or HGLM

Time to jump:
   Censored time to event data-> use survival analysis methods
   Kaplan-Meier
   Cox-proportional Hazards
   (Parametric methods)
   ‘Frailty’ models to allow for random effects
## GLM with Random effects

<table>
<thead>
<tr>
<th>GLM:</th>
<th>GLMM:</th>
<th>HGLM:</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E(y) = \mu$</td>
<td>$E(y</td>
<td>u) = \mu$</td>
</tr>
<tr>
<td>$\eta = G(\mu) = X\beta$</td>
<td>$\eta = G(\mu) = X\beta + Zu$</td>
<td>$\eta = G(\mu) = X\beta + Z\nu$</td>
</tr>
<tr>
<td>$V(y) = \phi V(m)$</td>
<td>$V(y</td>
<td>u) = \phi V(\mu)$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Commands:</th>
<th>GLM Commands</th>
<th>GLMM Commands</th>
<th>HGLM Commands</th>
</tr>
</thead>
<tbody>
<tr>
<td>MODEL</td>
<td>GLMM</td>
<td>HGRANDOM</td>
<td></td>
</tr>
<tr>
<td>FIT</td>
<td>IRREML</td>
<td>HGFIXED</td>
<td></td>
</tr>
</tbody>
</table>

$u \sim \text{Normal}$

$u \sim H(\lambda)$

---

The New Zealand Institute for Plant & Food Research Limited
Survival Analysis

Model distribution $f(t)$ of times to an event, e.g. $f(t) = \log\text{Normal}$

$F(t)$: Cumulative distribution

$S(t)$: Survivor function, $= 1 - F(t)$

$h(t)$: hazard function, prob('jump' at time $t$ / not 'jumped' before time $t$) $= f(t)/S(t)$

$H(t)$: Cumulative hazard, $= -\log(S(t))$
Proportional hazards:

For treatment N and baseline treatment 0
\[ h(t)_N = \phi h(t)_0 \]  \( \phi \) is the hazard ratio.

Individual with treatment N is \( \phi \) times as likely to ‘jump’ at any time compared to individual of treatment 0.

Put \( \phi = e^{\beta'X} \) to link to explanatory variables \( X \)

If \( X \) contains random components -> ‘Frailty Model’
Cox Proportional Hazards

Cox model:
*partial likelihood method*

- treat baseline hazard \( h_0 \) as block factor
- defined only where an event death occurs \( \rightarrow H_0 \) is step function, with jumps at each event time.

\( \beta \) can be estimated, without \( h_0(t) \) being estimated

Implementation as GLM

Log. Likelihood = Poisson Likelihood + a constant.

Involves linear predictor:

\[
\log(H(t)_i) = \log(H(t)_0) + \beta'X_i
\]

Estimate \( H(t)_0 \) as piece-wise exponential

Ref: Aitken *et al*
Implementation as GLM in GenStat

Expand data-set
1 row of data for each event time for each individual
Create Interval factor, with one level for each event time
Can do this with RPHVECTORS
Done implicitly within RPHFIT, RPROPORTIONAL

| Event times for 3 thrips: 3,5,1 |
|-----------------|-----------------|-----------------|-----------------|
| ThripsNo | Interval | Jump | Treat |
| 1      | 1        | 0    | C    |
| 1      | 3        | 1    | C    |
| 2      | 1        | 0    | Mi   |
| 2      | 3        | 0    | Mi   |
| 2      | 5        | 1    | Mi   |
| 3      | 1        | 1    | C    |

Non-Frailty:
MODEL [dist=Pois; link=log] Jump
FIT Interval+Treat

Frailty, with random thrips effect
HGRAND [dist=n; link=id] ThripsNo
HGFIX [dist=p; link=log] Interval+Treat
HGANAL Jump
Results: Percentage Jumped and Time Spent Moving

No thrips jumped with the control (air)
MI: Fewer thrips jumped with bean than with Chrysanth. or Pepper
Of those that jumped, a lower % of time was spent moving for Chrys. Leaf
Results: Time to jump
Kaplan-Meier Survivor Function

Only available for Mi thrips, as no control thrips ‘jumped’

Thrips on bean much slower to jump

Maybe thrips on Chrys. Flower slower than leaf or pepper after 50% have jumped
Cox Proportional hazard modeling

Estimates for model:

<table>
<thead>
<tr>
<th>Plant</th>
<th>Hazard ratio $\varphi$ relative to Bean</th>
<th>$\log(\varphi)$ (s.e.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chrys. Leaf</td>
<td>3.22</td>
<td>1.17 (0.49)</td>
</tr>
<tr>
<td>Chrys. Petal</td>
<td>4.05</td>
<td>1.40 (0.48)</td>
</tr>
<tr>
<td>Pepper</td>
<td>3.79</td>
<td>1.33 (0.48)</td>
</tr>
<tr>
<td>Overall</td>
<td>3.67</td>
<td>1.30 (0.44)</td>
</tr>
</tbody>
</table>

$h(t)_P = \varphi h(t)_B$  
$h(t)_B$ is baseline hazard function, for Bean

Hazard ratios $\varphi$ for Chrys. and Pepper significantly larger than 1 ($p<0.001$).

Ratios fairly similar for Chrys (Leaf, Petal) and Pepper ($p=0.519$).

At any time $t$, if a thrips has not jumped, it is more than 3x as likely to jump from Chrys or Pepper as from Bean.
Cox Proportional hazard modeling - with Frailty

Model with random effect for individual thrips failed to converge => individual thrips within a treatment not varying more than expected
Models with ‘Replicate’ or ‘Order within replicate’ converged
BUT in neither case were the random effects significant (HGRTEST)

Estimates for model with ‘Replicate’ random effect:

<table>
<thead>
<tr>
<th>Plant</th>
<th>Hazard ratio $\phi$ relative to Bean</th>
<th>$\log(\phi)$ (s.e.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chrys. Leaf</td>
<td>3.43</td>
<td>1.23 (0.49)</td>
</tr>
<tr>
<td>Chrys. Flower</td>
<td>4.31</td>
<td>1.46 (0.49)</td>
</tr>
<tr>
<td>Pepper</td>
<td>3.97</td>
<td>1.38 (0.49)</td>
</tr>
<tr>
<td>Overall</td>
<td>3.87</td>
<td>1.35 (0.44)</td>
</tr>
</tbody>
</table>

Conclusions as for model without set, but hazard ratios are all slightly increased.
Comments

GLM:
   An essential tool!!

Random effect models, GLMM/ HGLM:
   - Usability is improving, but still a way to go
   - Sometimes feel there little gain over using ANOVA+Transform 😊

Time to Event/ Survival analysis methods:
   More widely applicable than you might think


For upper Conf Limits for 0 estimates: