412th Test Wing

War-Winning Capabilities … On Time, On Cost

Title: Parametric TLE

Date: 9-11 April 2019

Author: James Brownlow

Office 812 TSS/ENTR

Phone: 661 277-4843

U.S. AIR FORCE

Preamble:

- Applying parametric survival models to analyze target location error (TLE) was the brainchild of Todd Remund, Greg Hutto, and Jeff Beekman (Beekster).
- This presentation details how the parametric survival model was adapted to TLE.
- An example is given- input file and source code in Python are available

Introduction

- Basic idea: precision target location, germane to navigation, weapon delivery and target tracking, can be related to distance from a target, as well as other possible explanatory factors (elevation and azimuth angles to the target, for example)
- Use the approach presented in Meeker and Escobar, *Statistical Reliability*, to use a generalized linear model based on log-link distribution function.
- Model fitting capabilities exist in JMP, as well as code written in Python and R

CEP

- CEP, CE10 and CE90 defined as
  - radius of a circle about the target that has the property that 50%, 10%, or 90% (respectively) of tle values are within a circle of radius ‘CEP’ and so forth
- Using parametric survival model, estimate CEP (or CE90, CE10) as functions of range R and associated covariates
- Additionally, we’d like confidence intervals for CE estimates, an RMS error estimate and 95/90 tolerance limits

Development

- TLE error estimates are similar to reliability of a component;
  - The cumulative probability of failure can be related to time in service
  - Similarly, cumulative probability of TLE can be related to range to the target
- Reliability (parametric) modeling provides a way to specify CEP (or CEwhatever) as a function of range to target.

Parametric Model

- Summary: generalized linear model
- $P(T \leq \mathrm{t_i} \,|\boldsymbol{X}) = \Phi(\frac{log(ti)-\mu_i}{\sigma_i})$, where
  $\boldsymbol{\mu_i} = \mathrm{b}_0 + \mathrm{b}_1\,\boldsymbol{X}_i$ and $\boldsymbol{\sigma_i} = \mathrm{b}_2 + \mathrm{b}_3\,\boldsymbol{X}_i$
- Percentile estimates, RMS, tolerance intervals all given by Φ
- **X** an *n* x *m* design matrix, columns contain covariates
- Details in Meeker and Escobar, *Statistical Reliability*

A few details

- **X** , design matrix, has *n* rows, *m* columns, X(i,j) = i-th observation, explanatory variable j
- In terms of matrix algebra:
  - $\boldsymbol{\mu} = \boldsymbol{X\beta} + \boldsymbol{\epsilon},$

  $\boldsymbol{\mu}$ *is n x 1,* $\boldsymbol{X}$ *is n x m,* $\boldsymbol{\beta}$ *is m x 1, and* $\boldsymbol{\epsilon}$ is n x 1

  (similarly for $\boldsymbol{\sigma}$)
- Estimates of $\boldsymbol{\beta}$ found by maximum likelihood
- Log link functions:

Log Error	Link Distribution
Log normal	normal
Extreme value	Weibull
log-logistic	logistic

Link Equations

- The key is the link between the distribution of the log of a random variable and the distribution of the random variable;
  - Y~ Weibull, then log(Y) ~ gumbel (EVS)
  - Y ~ Normal, then log(Y) ~ lognormal
  - Y ~ Logistic, then log(Y) ~ loglogistic
- Fit a generalized linear model to the log(Y)- in the ‘normal’ case:

$$P(\log(Y) < y) = \Phi_{ln}^{-1}\left(\frac{y-\mu_r}{\sigma_r}\right),$$

$\mu_r, and\ \sigma_r$ functions of covariates-

Estimation

- Estimation of parameters, β_0, β_1, and σ is based on the likelihood function

$$likelihood(\beta_0, \beta_1, \sigma) = \prod_{i=1}^{n} \frac{1}{\sigma} \phi(\frac{y_i - \mu_i}{\sigma})$$

  - $\boldsymbol{\mu_i} = \beta_0 + \beta_1 \boldsymbol{X}_i$ and $\boldsymbol{\sigma_i} = \beta_2 + \beta_3 \boldsymbol{X}_i$
- Maximum likelihood estimates of β's

Normal-lognormal case

For $\Phi_{\ln}$, the log normal distribution function,

$$\Pr(Y \le y) = F(y;\mu,\sigma) = F(y;\beta_0,\beta_1,\sigma) = \Phi_{\ln}((y-\mu)/\sigma)$$

The quantile function for this model is

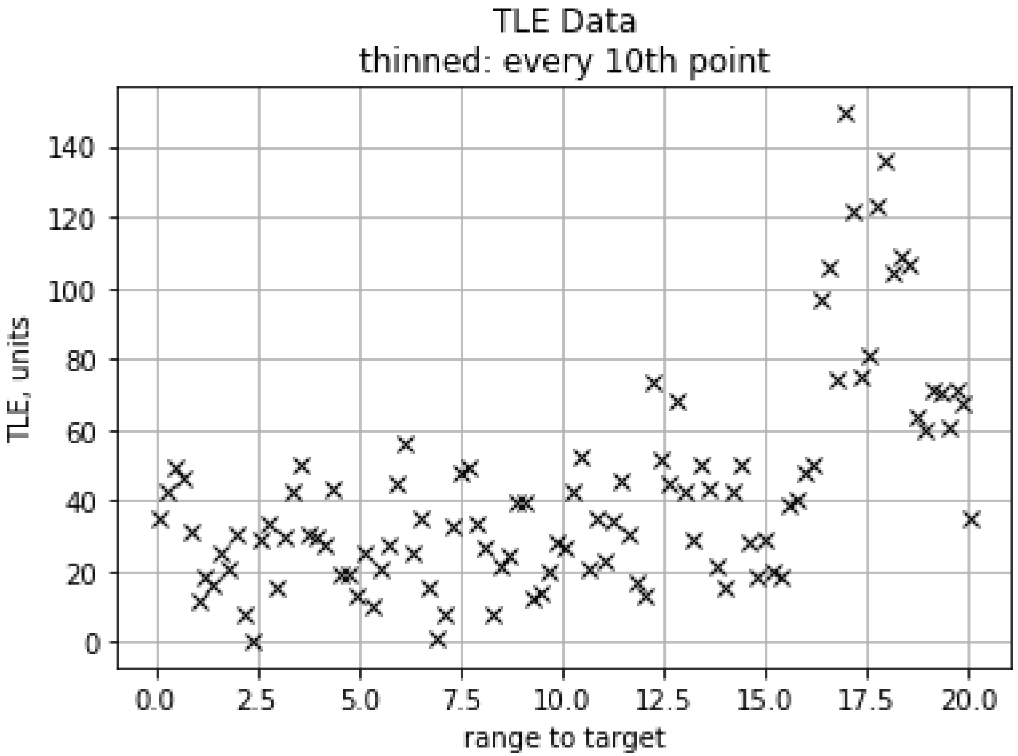
$$y_p(r) = \mu + \Phi_{\ln}{}^{-1}(p)\sigma = \beta_0 + \beta_1 r + \Phi_{\ln}{}^{-1}(p)\sigma$$

The quantile function then gives us probability curves for TLE, just select ‘p’

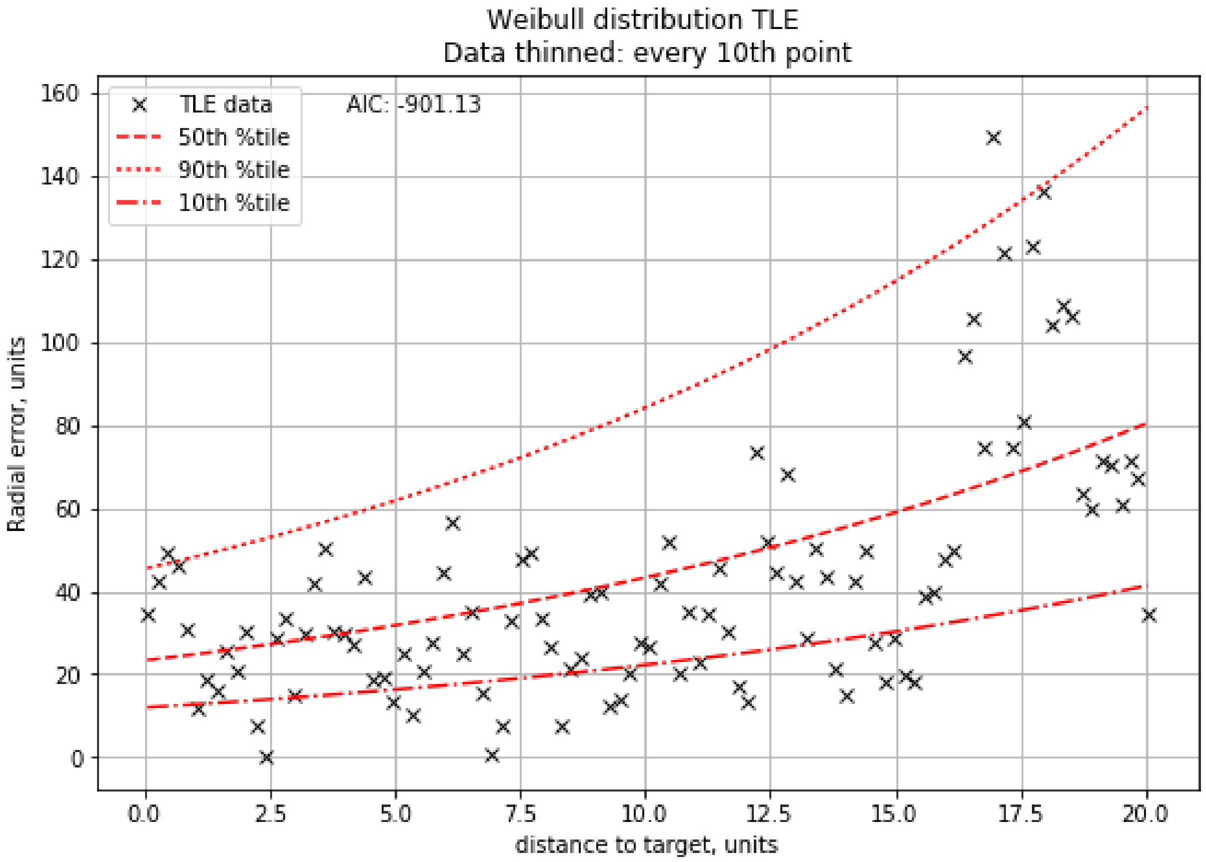
Upshot..

- Parametric modeling, if appropriate may avoid the effort to get IID errors
- Drawback- modeling assumes each realization (data run) is representative of a stationary, ergodic process
- Better approach may be a hierarchical model

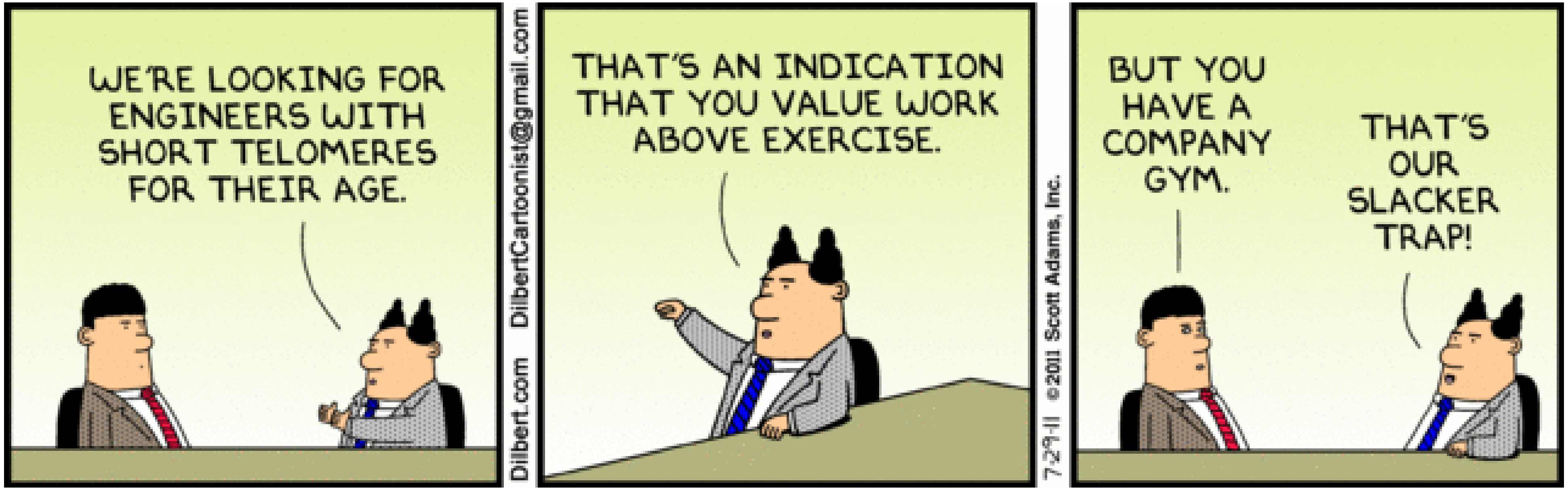
Example: Original data



Weibull distribution



Questions??



WE'RE LOOKING FOR ENGINEERS WITH SHORT TELOMERES FOR THEIR AGE.
THAT'S AN INDICATION THAT YOU VALUE WORK ABOVE EXERCISE.
BUT YOU HAVE A COMPANY GYM.
THAT'S OUR SLACKER TRAP!
Dilbert.com DilbertCartoonist@gmail.com
7-29-11 ©2011 Scott Adams, Inc.