Probabilistic Performance Assessment—Quick Takes
Alan Hutson
Department of Biostatistics and Bioinformatics

PPA model documents

- Fitting models to data
- Prediction
- Simulation
- Sensitivity Analysis
- Bayesian vs. Frequentist methods
- Spatial Smoothing
- My interpretation after reading several PPA documents
Toy Example-Fitting a Model

- \( C(t) = \beta e^{-\lambda t} \)
  - \( t \) = time
  - \( \beta \) = initial value at time \( t=0 \)
  - \( \lambda \) = rate parameter

- The function \( C(t) \) may be a “known” physical relationship or an approximation to a physical process

- The parameters \( \beta \) and \( \lambda \) may be known physical constants or estimated from the data

Toy Example-Known Parameters

\( C(t) \)

\( \beta = 0.1 \) and \( \lambda = 0.1 \), time was offset by 1990
Real Life Data-Measurement Error

Parameter Estimates

• More data → less uncertainty
  – Uncertainty can be quantified multiple ways
• Several estimation methods

- Assumption of the model's functional form is critical

<table>
<thead>
<tr>
<th>Estimate</th>
<th>Standard Error</th>
<th>t-Statistic</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>betah</td>
<td>0.111352</td>
<td>0.00475469</td>
<td>23.4194</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1.79915x10^-19</td>
</tr>
<tr>
<td>lambdah</td>
<td>0.106744</td>
<td>0.00699264</td>
<td>15.2652</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>8.41339x10^-15</td>
</tr>
</tbody>
</table>
Prediction

- \( \hat{C}(t) = \hat{\beta} e^{-\hat{\lambda}t} \)
- Say at the year \( t = 2025 \) we wish to predict \( C(2025) \)
- \( \hat{\beta} = 0.111352 \) and \( \hat{\lambda} = 0.106744 \)
- Predicted Value
  \[ \hat{C}(2025) = 0.00265556 \pm ? \]
  True value known only to God
  \( C(2025) = 0.00301974 \)

Prediction

- The relative error of the prediction depends upon the amount of data.
- There are many statistical/probabilistic ways to quantify this error including simulation methods
  – We could resample likely values for \( \beta \) and \( \lambda \) based on the data or an assumed distribution
- Again, also depends on the correct functional form for the model an accurate representation of a physical process
Toy Example—Poor Model/Prediction

- Assume incorrectly $C(t) = \beta + \lambda t$
- $t$ = time
- $\beta$ = initial value at time $t=0$
- $\lambda$ = rate parameter
- Nothing stops one from fitting the wrong model

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Standard Error</th>
<th>t-Statistic</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>betah</td>
<td>0.0751104</td>
<td>0.00447951</td>
<td>16.7676</td>
<td>$8.42987 \times 10^{-16}$</td>
</tr>
<tr>
<td>lambdah</td>
<td>-0.00286459</td>
<td>0.000274656</td>
<td>-10.4297</td>
<td>$5.71349 \times 10^{-11}$</td>
</tr>
</tbody>
</table>
Prediction

- \( \hat{C}(t) = \hat{\beta} + \hat{\lambda}t \)
- Say at time \( t = 2025 \)
- \( \hat{\beta} = 0.0751104 \) and \( \hat{\lambda} = -0.00286459 \)
- Predicted Value
  \( \hat{C}(2025) = -.0251502 \) +/- ?

Makes no physical sense.....

True value known only to God

\( C(2025) = .00301974 \)

Statistical Models

- Attributed to George Box

“All models are wrong, but some are useful”
Simulation

• Simulation is useful for complex models in terms of prediction error, behavior, etc

\[ \beta \]

• Resample values for \( \beta \) and \( \lambda \) to quantify the prediction error (or resample data)

Simulation

• Just like assuming the correct model there are several structural assumptions relative to the simulation of data
  – Correlation/dependence between the parameter distributions (temporal correlations)
  – Probability distribution for each parameter and/or joint distribution
    • Data based or assumed
    • Theoretical/large sample or physical properties
  – There are many simulation/resampling methodologies
Simulation

- There are theoretical approaches to “best” simulate data
- The goal is to approximate natural processes based on all of the available information to examine prediction error and the robustness of the assumed models

Simulation—Toy Example

- Let us assume our model assumptions were “accurate” or even “known”

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Standard Error</th>
<th>t-Statistic</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>betah</td>
<td>0.111352</td>
<td>0.00475469</td>
<td>23.4194</td>
<td>1.79915×10^{-19}</td>
</tr>
<tr>
<td>lambdah</td>
<td>0.106744</td>
<td>0.00699264</td>
<td>15.2652</td>
<td>8.41339×10^{-15}</td>
</tr>
</tbody>
</table>
Simulation—Toy Example

- We might assume the estimators for beta and lambda have normal probability distributions of this form (indep).

Simulation of Parameters

Example of 10 simulations (assume independent)

\[
\begin{array}{cc}
\beta & \lambda \\
0.143352 & 0.0588898 \\
0.124499 & 0.0094677 \\
0.102255 & 0.117203 \\
0.116917 & 0.117494 \\
0.14064 & 0.123542 \\
0.102595 & 0.165952 \\
0.0732821 & 0.141768 \\
0.0918887 & 0.121115 \\
0.124071 & 0.114804 \\
0.130925 & 0.107199 \\
\end{array}
\]
Simulation of Parameters

- From 100 simulations at time $t=2025$ we get 100 predicted $C(2025)$ values

![Histogram of Estimated $C(t)$](image)

- $C(2025)$ Mean=0.006 Standard Deviation=0.012 (biased estimate)
- Recall: True value is $C(2025)=.00301974$

Simulation of Parameters

- Where might we be “off-base”
- Likely picked the wrong probability distributions for $\beta$ and $\lambda$ estimates. For small samples in this scenario the distributions are likely to be asymmetrical
- Ignored the potential correlation structure between $\beta$ and $\lambda$ estimates
- We used estimates for $\beta$ and $\lambda$ ..random chance
Simulation of Parameters

- There are theoretical methodologies to examine the biases in the simulation methods.
- Other approaches to simulation, e.g. simulate the error term in the model, resample from the data, etc.
- The software used in the PPA modeling uses more deterministic resampling approaches, e.g. Latin Hypercube Resampling.

Sensitivity Analysis

- How sensitive is the model to the input parameters $\beta$ and $\lambda$?
- Not all parameters are “important”.
- Can statistically test if they are relevant given data.
- Can examine the sensitivity of each parameter.
  - E.g. what effect does a 10% change in either $\beta$ and $\lambda$ have on the predicted output.
Sensitivity Analysis

• Recall: \( \hat{C}(t) = \hat{\beta} e^{-\hat{\lambda} t} \)
• Say at time \( t = 2025 \)
• \( \hat{\beta} = 0.111352 \) and \( \hat{\lambda} = 0.106744 \)
• Predicted Value
  \( \hat{C}(2025) = 0.00265556 \)
  – What happens to \( \hat{C}(2025) \) if we increase
  \( \hat{\beta} = 0.111352 \) and \( \hat{\lambda} = 0.106744 \) by 10%

Sensitivity Analysis

– For \( \beta \) a 10% change leads to
  \( \hat{C}(2025) = 0.00292112, \) 10% change
– For \( \lambda \) a 10% change leads to
  \( \hat{C}(2025) = 0.00182769, \) 31% change

• The model is more “sensitive” to changes in \( \lambda \)

• The models considered for West Valley are
  much more complex than our simple example
  with several parameters and multiple
  interdependencies
Software Tools

• “The computer models GENII version 1.485 and LADTAP II were used to calculate site-specific unit dose factors (UDFs) for routine waterborne releases and dispersion of these effluents from the WVDP”

• “The RESRAD-BIOTA model was run using WVDP site-specific input concentrations of surface water, soil and sediment to output the annual dose to various categories of aquatic and terrestrial animals and plants”

Find some of the technical details

• http://resrad.evs.anl.gov/documents/
RESRAD

RESRAD Simulations
**RESRAD Simulations**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Parameter Type</th>
<th>Assigned Distribution Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>RESRAD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Density of contaminated zone (g/cm³)</td>
<td>P</td>
<td>Normal</td>
</tr>
<tr>
<td>Density of cover material (g/cm³)</td>
<td>P</td>
<td>Normal</td>
</tr>
<tr>
<td>Density of saturated zone (g/m³)</td>
<td>P</td>
<td>Normal</td>
</tr>
<tr>
<td>Depth of roots (m)</td>
<td>P</td>
<td>Uniform</td>
</tr>
<tr>
<td>Distribution coefficients (contaminated zone, unsaturated zones, and saturated zone)(cm³/g)</td>
<td>P</td>
<td>Lognormal</td>
</tr>
<tr>
<td>Saturated zone effective porosity</td>
<td>P</td>
<td>Normal</td>
</tr>
<tr>
<td>Saturated zone hydraulic conductivity (m/yr)</td>
<td>P</td>
<td>Lognormal</td>
</tr>
</tbody>
</table>

**RESRAD Simulations**

- Typical documented language, not necessarily specific to this project

Assignment of an appropriate distribution to a RESRAD or RESRAD-BUILD input parameter was determined primarily by the quantity of relevant data available. Documented distributions were used where available. However, data are often lacking for environmental exposure pathways. As fewer data became available, secondary types of information were used in conjunction with existing sample data in the distribution assignment task.
Fitting a Probability Distribution

• Several functional (“zillions”) forms to capture the shape and support (range) of a random variable. What to use?

• Can statistically test the “fit” of a distribution to the observed data

• Parametric distributions are typically defined by one to three parameters
Fitting a Probability Distribution

**Normal Distribution**

Probability density function

The red curve is the standard normal distribution.

---

Fitting a Probability Distribution

**Exponential**

Probability density function

---

17
Fitting a Probability Distribution

- The more data → more accurate parameter estimates
- Can pick the wrong distribution (can be tested)
- Several estimation methods
  - Maximum likelihood most well-known
  - Bayesian methods are also a likelihood based approach
  - i.e. given the data, what is the most likely values of the parameters that lead to that data
- One may consider “nonparametric” methods

Flag-WVDP Feb 2019 Inv Update

- Overutilization of truncated normal models
- Highly unlikely natural process
More realistic models

![Weibull (2-parameter) Probability density function](image)

More realistic models

![Log-normal Probability density function](image)
Fitting a Probability Distribution

• Several models to chose from
  – Truncated normal is very inflexible and has some key assumptions
  – There is no reason not to try to utilize the best fitting model or one that best approximates natural processes
  – Matter of equipoise—complicated multi-parameter models versus simpler lower dimension models.

Bayesian versus Frequentist

• Actually not that different. The key difference is a philosophical difference in how probability is defined
• Frequentist: Long run averages, e.g. flipping a coin millions of time will inform the probability of obtaining a “heads”
Bayesian versus Frequentist

• Bayesian: A person’s “belief” can be mapped to a probability space and updated with data, e.g. I believe the probability of “heads” is random around 0.5 with error. More data will “update” my belief of what the probability of heads is.
  – A noninformative prior would be no belief to start.....just the probability of heads is between 0 and 1

• The debate over which statistical approach is better has continued for decades.

Bayesian versus Frequentist

• Both rely on choice of correct model with functional forms

• Bayesian methods requires a prior distribution(s) about the model parameter(s)
  – Informative or non-informative
  – The prior is based on “belief”, not data
  – Data drives the posterior distribution
    • I am making an assumption that the PPA work is based on expert opinion and not a posterior distribution
Bayesian versus Frequentist

• Both approaches rely on data to provide accurate predictions
• In general, frequentist and Bayesian methods (should) yield similar results
Activity Distribution by Burial Record

- My guess is that they used a normal prior distribution and normal posterior distribution and fit the density model with one observation. Or this is just an illustration.
- Had to assume some scale value not based on data...seems to be some pooled value.
- Not exactly sure what these figures represent.
Total Inventory

• Not sure what is going on here.....certainly fitting non-normal distributions...Bayesian or frequentist.....probably nonparametric

Tank 8D-1 non-STS

• Unclear if this is a prior or posterior?? Is based on one data point or no data or much data
• Choice of scale seems to be pooled across waste forms
Brief

• Poor choice of models is generally not overcome by data (other approaches) for either frequentist or Bayesian parametric methods
• Bayesian methods tend to be anti-conservative, i.e. one can incorporate strong ‘beliefs’, higher likelihood of false positive results but can be more efficient if ‘correct’ prior distribution is chosen

Brief

• Frequentist methods are data driven only, conservative
• With the same ‘base’ distribution both tend to converge to the same answer as more data is gathered
• very similar mathematical properties when a non-informative prior is utilized.
**Brief**

- Reading the GENII version 1.485, LADTAP II RESRAD-BIOTA software documentation
  - Attempt to model a physical system
  - Uses ‘validated’ equations to model various interactive processes
  - Can modify model parameters
  - Simulations within the software are based on expert defined probability distributions, not data per se

**Spatial Smoothing**

- Estimate future and current distribution of X over a geographic region
- Interpolation versus extrapolation
- Covariance is a function of distance
- Evenly spaced grid of points for estimation
Flag-WVDP Feb 2019 Inv Update

Convex Hull-interpolation vs extrapolation