From last time:

Rep system, ROC(t)

$W(t) \rightarrow t_1 \rightarrow t_2 \rightarrow t_3 \rightarrow \text{Inter-event times}$

ROC (intensity) = $\lambda(t)$

(1) If we divide time into intervals, then for the process

$0 = h_1 < h_2 < \cdots$

$D_1 \text{ such that } D_2 \text{ such that } D_3 \text{ such that } D_i-1 \text{ such that } y_1 \text{ such that } y_2 \text{ such that } y_3 \text{ such that } y_i \text{ such that } h_i$

\[ \hat{W}(h_k) = \sum_{i=1}^{k} \frac{D_i}{y_i} \]

(2) Letting $h_1, h_2, \ldots$ more and more dense,

\[ \Rightarrow \hat{W}(t) = \sum_{t_i \leq t} \frac{D_i}{y_i(t_i)} \]

Variance of $\hat{W}$:

\[ \Delta (1) \text{ If the process is NHPP with } \lambda(t) \]

Then $D_i \sim \text{Poisson } \left( \lambda(t_i) \right)$

\[ \left( y_i, \Delta \left( \hat{W}(t_i) - \hat{W}(h_{i-1}) \right) \right) \]

Number of processes
We use that if \( X_1, X_2, \ldots, X_n \) are \( \text{Poisson}(\lambda_i) \) then \( \sum_{i=1}^{n} X_i \sim \text{Poisson}(\sum_{i=1}^{n} \lambda_i) \).

Then \( E(D_i) = y_i (W(k) - W(k_{i-1})) \)
and \( \text{Var}(D_i) = E(D_i) \) [property of Poisson & \( D_1, D_2, \ldots \) independent under property of NHRP].

Thus \( \text{Var}(W(k)) = \text{Var}\left( \sum_{i=1}^{k} D_i \right) \)

\[ = \sum_{i=1}^{k} \frac{1}{y_i^2} \text{Var}(D_i) \]

\[ = \sum_{i=1}^{k} \frac{1}{y_i^2} E(D_i) \]

But \( E(D_i) \) can be estimated by \( D_i \), so

\[ \hat{\text{Var}}(W(k)) = \sum_{i=1}^{k} \frac{\hat{D}_i}{y_i^2} \left( \frac{\hat{W}(k)}{k} = \sum_{i=1}^{k} \frac{D_i}{y_i} \right) \]

for an NHRP.

Taking limit: \( \text{Var}(W(\infty)) = \sum_{i=1}^{\infty} \frac{d(i)}{y(i)^2} \).
But — with not NPP: $W(t)$ still valid, but other $V(t)$ has to be used.

Slides 168–177 gives example etc. & 185–187 (value set data).

MINITAB: Stat > Reliability/Survival
  > Repairable System Analysis
  > Nonparametric Growth Curve
-4-

Parametric estimation in NHPP:

NHPP is characterized by the "intensity" \( \lambda(t) \).

**Popular models:**

\[
\begin{align*}
W(c,t) &= \int_0^c \lambda(u) \, du = E[N(c)] \\
W(s,t) &= \int_s^t \lambda(u) \, du = E[N(s,t)]
\end{align*}
\]

# events in interval \((s,t)\)

**Power law process NHPP:**

\[
\lambda(t) = \lambda \beta t^{\beta - 1} \quad \text{for } \beta > 0.
\]

\[
\begin{align*}
\downarrow & \quad \beta < 1 \\
\uparrow & \quad \beta > 1
\end{align*}
\]

**NHPP:** \( \beta = 1 \)

\[
W(t) = \int_0^t \lambda(u) \, du = \int_0^t \lambda_0 u^{\beta-1} \, du = \frac{t^\beta}{\beta}
\]

[Similar to Weibull intensity] \[ \sum \frac{\alpha}{\theta} \left( \frac{t}{\theta} \right)^{\alpha-1} \]

\[
W(s,t) = \Lambda(t^\beta - s^\beta)
\]
Log-linear MTPP:

\[ \omega(t) = e^{\alpha + \beta t} \]

- If \( \beta < 0 \)
- If \( \beta > 0 \)
- \( \beta = 0 \)

Let \( W(t) = \int_0^t e^{\alpha + \beta u} du \)

\[ = e^\alpha \int_0^t e^{\beta u} du \]
\[ = e^\alpha \left[ \frac{1}{\beta} e^{\beta u} \right]_0^t \]
\[ = \frac{e^\alpha}{\beta} (e^{\beta t} - 1) \]
\[ W(s,t) = \frac{e^\alpha}{\beta} (e^{\beta t} - e^{\beta s}) \]
Likelihood function for NHPP data

Suppose we observe one process; record \( w(t) \)

\[
0 < s_1 < s_2 < \ldots < s_N < T
\]

- Censoring

Observe interval \([0, T]\) fixed, given

# events \( N \) is random variable

\( N \sim \text{Poisson}(W(t)) \)

Times of events: \( 0 < s_1 < s_2 < \ldots < s_N < T \)

Parametric model: Write \( \rho \) as \( w(t; \theta) \)

\( e.g., w(t; \lambda, \beta) = \lambda \beta t^{\beta-1} \) for powerlaw

Define also \( W(t; \theta) = \int_0^t w(u; \theta) du = E[N(t)] \)

\[ W(s, t; \theta) = \int_s^t w(u; \theta) du = E[N(s, t)] \]
Divide line axis at \( h_0 = 0 < h_1 < h_2 < \ldots < h_r = \infty \)

\[ D_i = \# \text{events in } (h_{i-1}, h_i) \]

**Observations** are \( D_1, D_2, \ldots, D_r \)

which are independent (property of Poisson) and Poisson distributed.

Hence likelihood is

\[ L(\Theta) = P(D_1 = d_1, D_2 = d_2, \ldots, D_r = d_r) \]

\[ = \prod_{i=1}^{r} P(D_i = d_i) = \prod_{i=1}^{r} \frac{W(h_{i-1}, h_i; \Theta)}{d_i!} e^{-W(h_{i-1}, h_i)} \]

\[ \left\{ \prod_{i=1}^{r} \frac{W(h_{i-1}, h_i; \Theta)}{d_i!} \right\} \cdot e^{-\sum_{i=1}^{r} W(h_{i-1}, h_i)} \]

But

\[ \sum_{i=1}^{r} W(h_{i-1}, h_i; \Theta) = \sum_{i=1}^{r} \int_{h_{i-1}}^{h_i} w(u; \Theta) \, du \]

\[ = \int_{0}^{\infty} w(u; \Theta) \, du = W(\infty; \Theta) \]
\[ L(\theta) = \prod_{i=1}^{n} \frac{W(h_{i-1}, h_{i}, \theta)}{d_{i}} \]  

If data are given in groups like this, then we use this to find MLE.

If times are given exactly as \( s_{1}, s_{2}, \ldots, s_{n} \), then we let the grid of \( h_{i} \) be more and more dense and get in the limit \( 0 \) or \( 1 \) event in each interval \( (h_{i-1}, h_{i}) \).

Now when \( d_{i} = 0 \) is the contribution to the product \( W(h_{i-1}, h_{i}, \theta) \)  

\[ \prod_{i=1}^{n} \frac{W(h_{i-1}, h_{i}, \theta)}{d_{i}} = 1 \]

so we can ignore it in \( L(\theta) \).

When \( d_{i} = 1 \) is the contribution

\[ W(h_{i-1}, h_{i}, \theta) = \int_{h_{i-1}}^{h_{i}} w(u; \theta) du = \int_{h_{i-1}}^{h_{i}} \frac{w(s_{k})}{h_{i} - h_{i-1}} du \]

\[ \times w(s_{k})(h_{i} - h_{i-1}) \]

the single \( s_{k} \) in the interval
and since the hi-thi are given by us, we let the contribution to the likelihood be \( w(s_i^j; \theta) \).

Here

\[
L(\theta) = \prod_{i=1}^{N} w(s_i^j; \theta) e^{-W(\tau; \theta)}
\]

Similar to the \( \prod_{i=1}^{n} f(t_i^j; \theta) \) when we have \( n \) i.i.d. observed failures \( t_1, t_2, \ldots, t_n \).

See also slides 191-192 (taken from another course).

Note \( \bigcirc \) is for a single system.

If we have data for several systems with the same \( w(t_i^j; \theta) \):

\[
\bigcirc \quad s_{ij} \quad s_{ij} \quad \cdots \quad s_{ij} \quad \tau_j
\]

\( j = 1, 2, \ldots m \).
The likelihood is the product:

\[ L(\theta) = \prod_{j=1}^{m} \prod_{i=1}^{N_j} w(s_{ij}; \theta) e^{-W(c_j; \theta)} \]

or log-likelihood is

\[ l(\theta) = \sum_{j=1}^{m} \left[ \sum_{i=1}^{N_j} \ln w(s_{ij}; \theta) - W(c_j; \theta) \right] \]

**Example:**

\[ w(t; \lambda, \beta) = \frac{1}{\beta} t^{\beta - 1} \]
\[ W(t; \lambda, \beta) = \frac{1}{\beta} t^\beta \]

\[ \ln w(t; \lambda, \beta) = \ln t + \ln \beta + (\beta - 1) \ln t \]

\[ l(\theta; \beta) = \sum_{j=1}^{m} \left[ \sum_{i=1}^{N_j} (\ln t + \ln \beta + (\beta - 1) \ln s_{ij}) - \frac{c_j^\beta}{\beta} \right] \]
\[
\sum_{j=1}^{m} \sum_{i=1}^{N_j} \ln \lambda + \sum_{j=1}^{m} \sum_{i=1}^{N_j} \ln \beta_i + \sum_{j=1}^{m} \sum_{i=1}^{N_j} (\beta - 1) \ln S_{ij} - \sum_{j=1}^{m} \beta_i \\
\]

\[
= N \ln \lambda + m \ln \beta + (\beta - 1) S - \sum_{j=1}^{m} \beta_i \\
\]

Put \(N = \sum_{j=1}^{m} N_j\) where \(S = \sum_{j=1}^{m} \sum_{i=1}^{N_j} \ln S_{ij}\) to find the MLEs:

\[
\frac{\partial l}{\partial \lambda} = \frac{N}{\lambda} - \sum_{j=1}^{m} \beta_i = 0 \quad (1)
\]

\[
\frac{\partial l}{\partial \beta} = \frac{N}{\beta} + S - \lambda \sum_{j=1}^{m} (\beta_i - 1) \ln \beta_i = 0 \quad \text{for}\ j = 1, \ldots, m
\]

(1) gives \(\lambda = \frac{N}{\sum_{j=1}^{m} \beta_i}\)
Put into (2):

\[
\frac{N}{\beta} + S' - \frac{N \sum \frac{\hat{\epsilon}_j}{\beta} \ln \frac{\hat{\epsilon}_j}{\beta}}{\sum \frac{\hat{\epsilon}_j}{\beta}} = 0
\]

Can solve this numerically for \( \beta \) to get \( \hat{\beta} \)
and then put \( \hat{S} = N/\sum \frac{\hat{\epsilon}_j}{\beta} \)

Special cases:

If all \( \hat{\epsilon}_j \equiv \epsilon \) are equal:

\[
\frac{N}{\beta} + S' - \frac{Nm \epsilon \ln \frac{\epsilon}{\beta}}{m \epsilon} = 0
\]

So:

\[
\hat{\beta} = \frac{N}{\beta} + S - N \ln \frac{\epsilon}{\beta} = 0
\]

\[
\frac{N}{\beta} = N \ln \frac{\epsilon}{\beta} - S
\]

\[
\hat{\beta} = \frac{N}{N \ln \frac{\epsilon}{\beta} - S}
\]

So in this case

\[
\hat{\beta} = \frac{N}{N \ln \frac{\epsilon}{\beta} - S}
\]

\[\hat{\beta} \hat{S} \]
With profile likelihood:

If \( \beta \) is known, then we have

\[
\hat{\lambda}(\beta) = \frac{N}{\sum_{j=1}^{m} \tau_j^\beta}
\]

Profile likelihood of \( \beta \) is therefore

\[
\tilde{l}(\beta) = l(\hat{\lambda}(\beta), \beta) = \]

\[N \ln \hat{\lambda}(\beta) + N \ln \beta + (\beta - 1) S - \hat{\lambda}(\beta) \cdot \sum_{j=1}^{m} \tau_j^\beta \]

\[= N \ln N - N \ln \left( \sum_{j=1}^{m} \tau_j^\beta \right) + N \ln \beta + (\beta - 1) S - N\]

Then (1) Maximize \( \tilde{l}(\beta) \) to find \( \hat{\beta} \)

(2) Put \( \hat{\lambda} = \hat{\lambda}(\hat{\beta}) \).
In simple example:

\[ \tilde{\ell}(\beta) = 6 \ln 6 - 6 \ln (20^\beta + 30^\beta + 10^\beta) \]

\[ + 6 \ln \beta + \ln (\beta - 1) \approx 13.6466 - 6 \]

\[ \approx 5 \]

**Note:** (Slide 193)

Read off \( \hat{\beta} = 1.20 \)

\((\hat{\beta} = 1.19423 \text{ Minitab})\)

\[ \lambda (\hat{\beta}) = \frac{6}{20^\beta + 30^\beta + 10^\beta} \]

\[ \Rightarrow \lambda = 0.0538 \quad (\hat{\beta} = 1.20) \]

\[ 0.0598 \quad (\hat{\beta} = 1.19423) \]
Confidence interval for $\beta$ using likelihood theory:

\[ W(\beta) = 2 \left( l(\hat{\alpha}(\beta), \beta) - l(\hat{\alpha}(\beta), \beta) \right) \]

\[ = -l(\hat{\alpha}, \beta) \]

\[ \propto \chi^2_1 \]

\[ \Rightarrow \text{(as before) that a 95% (approx) conf. interval for } \beta \text{ is obtained by cutting off } \text{profile log-likelihood at maximum value} \]

\[ -1.92, \text{ i.e.} \]

\[ -19.71 - 1.92 = -21.63 \]

Read off from slide 192: Approx 95%: (0.50, 2.25)
Observed information etc.

Recall

\[ l(\lambda, \beta) = N \ln \lambda + N \ln \beta + (\beta - 1) S - \frac{1}{\beta} \sum_{j=1}^{m} \xi_j^\beta \]

where \( N = \sum_{j=1}^{m} N_j \)
\( S = \sum_{j=1}^{m} \sum_{c=1}^{\xi_j} \ln \xi_j \)

\[ \frac{\partial l}{\partial \lambda} = \frac{N}{\lambda} - \sum_{j=1}^{m} \xi_j^\beta \]

\[ \frac{\partial^2 l}{\partial \lambda^2} = -\frac{N}{\lambda^2} \]

\[ \frac{\partial^2 l}{\partial \lambda \partial \beta} = -\sum_{j=1}^{m} (\ln \xi_j) \xi_j^\beta \]

\[ \frac{\partial l}{\partial \beta} = \frac{N}{\beta} + S - \frac{1}{\beta} \sum_{j=1}^{m} (\ln \xi_j) \xi_j^\beta \]

\[ \frac{\partial^2 l}{\partial \beta^2} = -\frac{N}{\beta^2} - \frac{1}{\beta} \sum_{j=1}^{m} (\ln \xi_j)^2 \xi_j^\beta \]
Information matrix: Observed Information Matrix

\[
\begin{bmatrix}
- \frac{\partial^2 I}{\partial \alpha^2} & - \frac{\partial^2 I}{\partial \alpha \partial \beta} \\
- \frac{\partial^2 I}{\partial \alpha \partial \beta} & - \frac{\partial^2 I}{\partial \beta^2}
\end{bmatrix}
\]

\[\lambda = \frac{1}{\alpha}, \quad \beta = \beta\]

\[w = \sqrt{\frac{1}{\lambda}} \left( \sum \hat{\epsilon}_i \beta (\ln \hat{y}_i) \right)
\]

\[\sum \hat{\epsilon}_i \beta (\ln \hat{y}_i) + \frac{w}{\beta^2} + \frac{1}{\lambda} \left( \sum \hat{\epsilon}_i \beta (\ln \hat{y}_i)^2 \right)
\]

**Example:** In simple example:

\[
\begin{bmatrix}
\frac{6}{0.0538^2} = 2072.9 \\
347.03 \\
\frac{6}{1.2^2} + 0.0538 \cdot 1096 = 62.1015
\end{bmatrix}
\]

\[
\begin{bmatrix}
0.0060 - 0.0332 \\
0.1984
\end{bmatrix}
\]
so that \( Var(\beta) = 0.1984 \)
\[ SD(\beta) = \sqrt{0.1984} = 0.4454 \]

Standard interval 95%:
\[ \beta \pm 1.96 \cdot 0.4454 \]
\[ \begin{align*}
1.2 & \end{align*} \]

= \[ [0.327, 2.073] \] This is the interval MINITAB gives here

but MINITAB should have used the "standard interval for positive parameters" i.e.
\[ \pm 1.96 \frac{0.4454}{1.2} \]
\[ \begin{align*}
1.2 & \end{align*} \]

i.e. \[ [0.58, 2.48] \]

(which is also closer to the likelihood interval).
\[ \text{Var} \alpha = 0.0060 \]
\[ \Rightarrow \text{SD}(\alpha) = 0.0775 \]

Minitab output:
\[ \theta \text{ when } \theta = \frac{1}{8} \]
\[ \theta = \{ \theta \} \]

Standard 95%:
\[ 1 \pm 1.96 \cdot 0.0775 \]
\[ 0.0548 \]
\[ (0, 0.2062) \]

\[ \text{while standard for positive:} \]
\[ \frac{1}{e} \pm 1.96 \cdot \frac{0.0775}{\sqrt{e}} \]
\[ 0.0548 \cdot e \]
\[ (0.0034, 0.8301) \]

Note:
\[ \rho(\alpha, \beta) = \frac{\text{Cov}(\alpha, \beta)}{\sqrt{\text{Var}\alpha \cdot \text{Var}\beta}} = -0.0332 \]

Correlation coefficient:
\[ -0.9623 \]

Very strong correlation.

This is due to a functional relationship.
\[
\frac{y^N}{z} = \sum \bar{z} \cdot \beta
\]
\[
\frac{\alpha}{z} = \frac{N}{\sum \bar{z} \cdot \beta} = \frac{6}{30\beta + 20\beta + 10\beta}
\]

See Slide 194

(CAN DO CORRESPONDING LIKELIHOOD ANALYSIS FOR LOG-LINEAR NHPP)

MINITAB vs. "MINE" POWER LAW ESTIMATION:

MINITAB: \[ A(t) = \frac{\text{Shape}}{\text{Scale}} \left( \frac{t}{\text{Scale}} \right)^\text{Shape-1} \]

MINE: Power law: \[ A(t) = \frac{1}{\beta} t^{\beta-1} \]

\[ \Rightarrow \text{shape} = \beta \quad (\text{OK}) \]

\[ \lambda = \frac{\text{Shape}}{\text{Scale}} = \frac{1.202}{11.3803} = 0.106 \]

From MINITAB
Example: Campus data SLIDE 195-196.

Single system:

Power law: \( w(t) = \lambda \beta t^{\beta-1} \), gives

\[ \beta = 1.22, \quad \lambda = \frac{1}{0.553^{1.22}} = 2.06 \]

Loglinear: \( w(t) = e^{x + \beta t} \)

\[ \lambda = 1.01, \quad \beta = 0.0373 \]

Slide 195 shows:

\[ W(t): \] Nelson-Aalen

Power law: \( \hat{W}(t) = \hat{A} t^{\hat{\beta}} = 2.06, t^{1.22} \)

Log-linear: \( \hat{W}(t) = \frac{e^x}{\hat{\beta}} (e^{\hat{\beta} t} - 1) \)