Smooth Goodness-of-Fit Tests in Hazard-Based Models

by

Edsel A. Peña
Department of Statistics
University of South Carolina at Columbia
E-Mail: pena@stat.sc.edu

Univ. of Georgia Colloquium Talk
Department of Statistics
November 16, 2000

Research supported by an NIH Grant

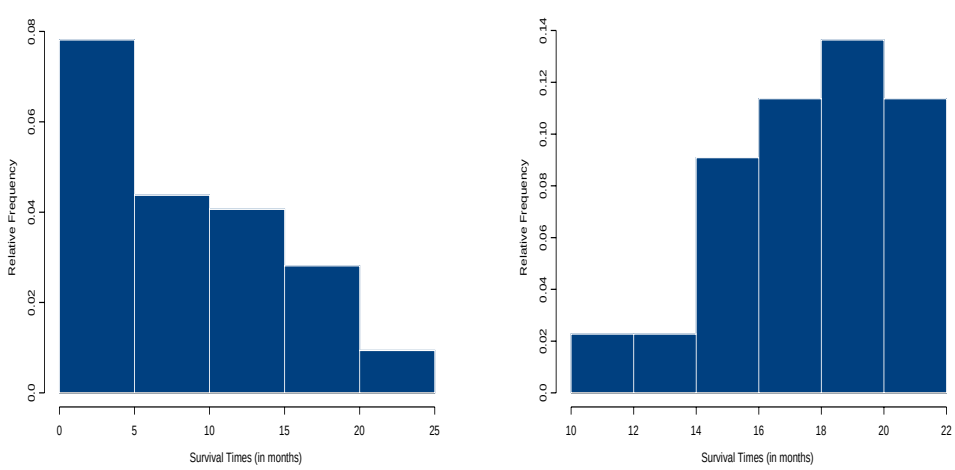
1. Practical Problem

- Right-censored survival data for lung cancer patients from Gatsonis, Hsieh and Korwar (1985).

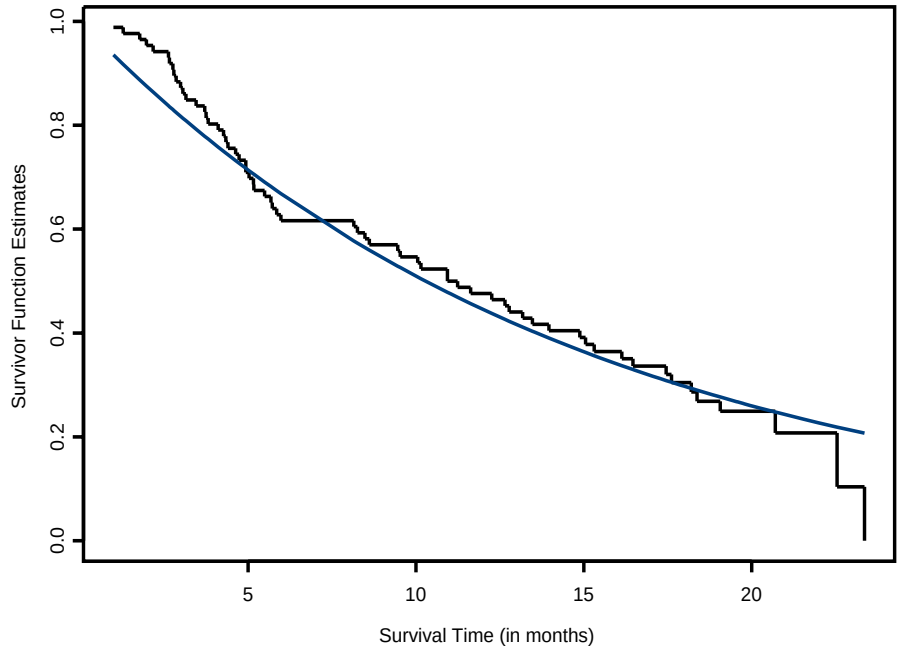
- Survival times (in months): $n = 86$ with 23 right-censored.

  0.99, 1.28, 1.77, 1.97, 2.17, 2.63, 2.66, 2.76, 2.79, 2.86, 2.99, 3.06, 3.15, 3.45, 3.71, 3.75, 3.81, 4.11, 4.27, 4.34, 4.40, 4.63, 4.73, 4.93, 4.93, 5.03, 5.16, 5.17, 5.49, 5.68, 5.72, 5.85, 5.98, 8.15, 8.26, 8.48, 8.61, 9.46, 9.53, 10.05, 10.15, 10.94, 10.94, 11.04+, 11.24, 11.63, 12.26, 12.65, 12.78, 13.18, 13.47, 13.53+, 13.96, 14.23+, 14.65+, 14.88, 14.91+, 15.05, 15.31, 15.47+, 16.13, 16.46, 16.49+, 17.05+, 17.28+, 17.45, 17.61, 17.88+, 17.97+, 18.20, 18.37 18.83+ 19.06 19.55+ 19.58+ 19.75+ 19.78+ 19.95+ 20.04+, 20.24+, 20.70, 20.73+, 21.55+, 21.98+, 22.54, 23.36

- Probability histograms of the Complete and Censored Values.



- Product-limit estimator and best-fitting exponential survivor function.

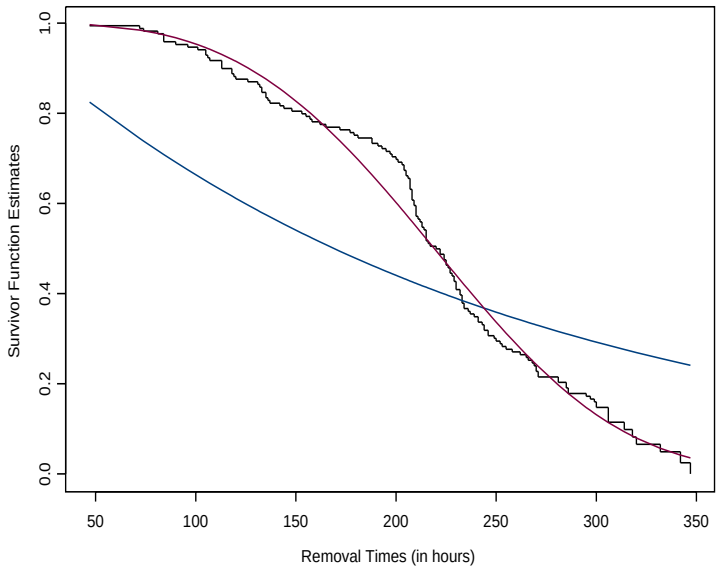


- **Question:** Did the survival data come from the family of exponential distributions? Or was it from the family of Weibull distributions?

- **Another Data Set:** Times to withdrawal (in hours) of 171 car tires in Davis and Lawrance (1989, *Scand. J. Statist.*) from a car-tire testing study, with withdrawal either due to failure or right-censoring.

- The pneumatic tires were subjected to laboratory testing by rotating each tire against a steel drum until either failure [actually, there were several competing causes] or removal (right-censoring).

- Plots of the product-limit estimator, best-fitting exponential, and best-fitting Weibull distribution.



- **Question:** Is this failure-time car tires data consistent with an underlying Weibull distribution?

- Both situations point to a **goodness-of-fit** problem with right-censored data.

2. Densities and Hazards

- T = a positive-valued continuous failure-time variable, e.g.,
  - ◇ time-to-failure of a mechanical or electronic system
  - ◇ time-to-occurrence of an event
  - ◇ survival time of a patient in a clinical trial

- $f(t)$ = density function of T.

$$f(t)\Delta t \approx \mathbf{P}\left\{T \in [t, t+\Delta t)\right\}.$$

- $F(t) = \mathbf{P}\{T \leq t\}$ = distribution function
- $\bar{F}(t) = 1 - F(t)$ = survivor function

- $\lambda(t) = f(t)/\bar{F}(t)$ = hazard rate function.

$$\lambda(t)\Delta t \approx \mathbf{P}\left\{T \in [t, t+\Delta t) | T \geq t\right\}.$$

- $\Lambda(t) = \int_0^t \lambda(w)\mathrm{d}w = -\log[\bar{F}(t)]$ = (cumulative) hazard function

- Equivalences:

$$\bar{F}(t) = \exp\{-\Lambda(t)\}$$

$$f(t) = \lambda(t)\exp\{-\Lambda(t)\}$$

- Two Common Examples:

  ◇ Exponential:

$$f(t;\eta) = \eta \exp\{-\eta t\}$$

$$\bar{F}(t;\eta) = \exp\{-\eta t\}$$

$$\lambda(t;\eta) = \eta$$

$$\Lambda(t;\eta) = \eta t$$

  ◇ Two-Parameter Weibull:

$$f(t;\alpha,\eta) = (\alpha\eta)(\eta t)^{\alpha-1} \exp\{-(\eta t)^{\alpha}\}$$

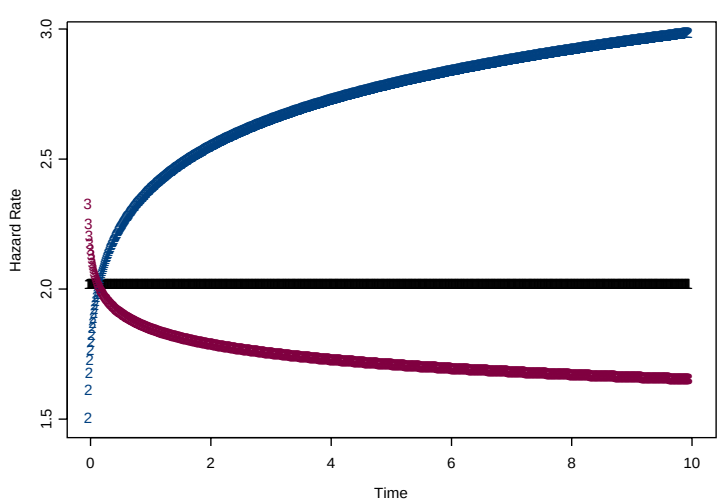
$$\bar{F}(t;\alpha,\eta) = \exp\{-(\eta t)^{\alpha}\}$$

$$\lambda(t;\alpha,\eta) = (\alpha\eta)(\eta t)^{\alpha-1}$$

$$\Lambda(t;\alpha,\eta) = (\eta t)^{\alpha}$$

  ◇ Qualitative Aspects from Plots of Hazards

Weibull Hazard Plots



3. A Pitch for Hazard-Based Modeling

- Advantages of specifying models via hazards:
  - ◇ Vantage point in density modeling: Time origin.
  - ◇ Vantage point in hazard modeling: 'Present, together with information that accumulated in the past.'
  - ◇ Qualitative aspects (e.g., IFR, or bath-tub) can be incorporated.
  - ◇ Facilitates incorporation of dynamic evolution of system or component.
  - ◇ Likelihood construction natural via product integrals.
  - ◇ Adapts well in the presence of right-censored or truncated data.
  - ◇ Conducive to modeling with point processes (popularized by Aalen; Andersen and Gill; etc.).

- Theory to be presented applicable to more general models, but will only consider the following hazard-based models.

  - ◇ **IID Model:** $T_1, T_2, \ldots, T_n$ IID with common unknown hazard rate function $\lambda(t)$. Observable vectors:

    $$(Z_1, \delta_1), (Z_2, \delta_2), \ldots, (Z_n, \delta_n)$$

    with

    $$\delta_i = 1 \Rightarrow T_i = Z_i$$

    $$\delta_i = 0 \Rightarrow T_i > Z_i.$$

  - ◇ **Cox PH Model** (also Andersen and Gill Model):

    $$(T_1, X_1), (T_2, X_2), \ldots, (T_n, X_n)$$

    with

    $$\lambda_{T|X}(t|X) = \lambda(t) \exp\{\beta^{\mathrm{t}} X\}$$

    $\lambda(\cdot)$ an unknown hazard rate function, and β a regression coefficient vector. Observable vectors:

    $$(Z_1, \delta_1, X_1), (Z_2, \delta_2, X_2), \ldots, (Z_n, \delta_n, X_n)$$

    with

    $$\delta_i = 1 \Rightarrow T_i = Z_i$$

    $$\delta_i = 0 \Rightarrow T_i > Z_i.$$

4. Problems and Issues

- **Problem (Goodness-of-Fit):** Given

$$\{(Z_i, \delta_i), i = 1, 2, \ldots, n\}$$

in the IID model, or

$$\{(Z_i, \delta_i, X_i), i = 1, 2, \ldots, n\}$$

in the Cox model, decide whether

$$\lambda(\cdot) \in \mathcal{C} = \{\lambda_0(\cdot; \eta) : \eta \in \Upsilon\}$$

where η is a nuisance parameter vector.

  - ◇ $\mathcal{C}$ could be: Exponential, Weibull, Gamma; or IFRA class.

  - ◇ Importance?

  - ◇ Previous works on GOF problem with censored data: Akritas (88, JASA), Hjort (90, AS), Hollander and Peña (92, JASA), Li and Doss (93, AS), and others.

  - ◇ Generalize Pearson's:

$$\chi_P^2 = \sum \frac{(O_j - \hat{E}_j)^2}{\hat{E}_j}?$$

  - ◇ Difficulty: O_j's not computable with censored data. Also, any optimality properties of existing tests?

- **Problem (Model Validation):** Given $\{(Z_i, \delta_i), i = 1, 2, \ldots, n\}$ or $\{(Z_i, \delta_i, X_i), i = 1, 2, \ldots, n\}$, how to assess model assumptions (e.g., IID assumption, model structure)?

  ◇ Unit Exponentiality Property (UEP)

  $$T \sim \Lambda(\cdot) \Rightarrow \Lambda(T) \sim \text{EXP}(1)$$

  ◇ IID model: If $\Lambda_0(\cdot)$ is the true hazard function, then with $R_i^0 = \Lambda_0(Z_i)$,

  $$(R_1^0, \delta_1), (R_2^0, \delta_2), \ldots, (R_n^0, \delta_n)$$

  is a right-censored sample from EXP(1).

  ◇ Since $\Lambda_0(\cdot)$ is not known, the R_i^0's are estimated by R_i's with

  $$R_i = \hat{\Lambda}(Z_i), \quad i = 1, 2, \ldots, n,$$

  $\hat{\Lambda}(\cdot)$ is an estimator of $\Lambda_0(\cdot)$ based on the (Z_i, δ_i)'s.

  ◇ **Idea:** (R_i, δ_i)'s *assumed to form an* **approximate** *right-censored sample from EXP(1),* so to validate model, test whether (R_i, δ_i)'s is a right-censored sample from EXP(1).

  ◇ **Question:** How good is this *approximation*, even in the limit???

◇ For Cox PH model, the analogous expressions for R_i^0 and R_i are:

$$R_i^0 = \Lambda_0(Z_i)\exp\{\beta X_i\}, \quad i = 1, 2, \ldots, n;$$

$$R_i = \hat{\Lambda}(Z_i)\exp\{\hat{\beta} X_i\}, \quad i = 1, 2, \ldots, n.$$

∗ $\hat{\Lambda}(\cdot)$ is an estimator of $\Lambda_0(\cdot)$ [e.g., Aalen-Breslow estimator].

∗ $\hat{\beta}$ is an estimator of β [partial likelihood MLE].

◇ R_i^0's are **true generalized residuals** (Cox and Snell (68, JRSS)).

◇ R_i's are **estimated generalized residuals.**

◇ Generalized residuals are analogs of linear model residuals:

"**(Observed Value)** minus **(Fitted Value)**."

◇ **Question:** Effects of substituting estimators for the unknown parameters, even when the estimators are consistent??

5. Class of Smooth GOF Tests

- Convert observed data

$$(Z_1, \delta_1, X_1), (Z_2, \delta_2, X_2), \ldots, (Z_n, \delta_n, X_n)$$

into stochastic processes.

- For $i = 1, 2, \ldots, n$, and $t \geq 0$, let

$$N_i(t) = I\{Z_i \leq t, \delta_i = 1\};$$

$$Y_i(t) = I\{Z_i \geq t\}.$$

-

$$A_i(t; \lambda(\cdot), \beta) = \int_0^t Y_i(w)\lambda(w) \exp\{\beta^{\mathrm{t}} X_i\} \mathrm{d}w;$$

$$M_i(t; \lambda(\cdot), \beta) = N_i(t) - A_i(t; \lambda(\cdot), \beta).$$

- If $\lambda_0(\cdot)$ and β_0 are the *true* parameters,

$$M^0(t) = (M_1(t; \lambda_0(\cdot), \beta_0), \ldots, M_n(t; \lambda_0(\cdot), \beta_0))$$

are orthogonal sq-int zero-mean martingales with predictable quadratic variation processes

$$\langle M_i^0, M_i^0 \rangle(t) = A_i(t; \lambda_0(\cdot), \beta_0).$$

- **Problem:** Test

$$H_0 : \lambda(\cdot) \in \mathcal{C} = \{\lambda_0(\cdot; \eta) : \eta \in \Upsilon\} \quad \text{versus} \quad H_1 : \lambda(\cdot) \notin \mathcal{C}.$$

- **Idea:** If $\lambda_0(\cdot)$ is the true hazard rate function, then under H_0 there is some $\eta_0 \in \Upsilon$ such that

$$\lambda_0(\cdot) = \lambda_0(\cdot; \eta_0).$$

- Define

$$\kappa(t; \eta) = \log\left[\frac{\lambda_0(t)}{\lambda_0(t; \eta)}\right].$$

- $\mathcal{K}$ = collection of such $\{\kappa(\cdot; \eta) : \eta \in \Upsilon\}$.

- Basis set (e.g., trigonometric, polynomial, wavelet, etc.) for $\mathcal{K}$:

$$\mathbf{\Psi} = \{\psi_1(\cdot; \eta), \psi_2(\cdot; \eta), \ldots\};$$

$$\kappa(t; \eta) = \sum_{j=1}^{\infty} \theta_j \psi(t; \eta) = \theta' \mathbf{\Psi}.$$

- For smoothing order K, approximate $\kappa(\cdot; \eta)$ by

$$\kappa(t; \eta) = \sum_{j=1}^{K} \theta_j \psi(t; \eta) = \theta_K' \mathbf{\Psi}_K.$$

- Equivalently,

$$\lambda_0(t) \approx \lambda_0(t; \eta) \exp\left\{\sum_{j=1}^{K} \theta_j \psi(t; \eta)\right\}.$$

- Define:

$$\mathcal{C}_K = \left\{\lambda_K(\cdot; \theta, \eta) = \lambda_0(\cdot; \eta) \exp\left\{\sum_{j=1}^{K} \theta_j \psi_j(\cdot; \eta)\right\} : \quad \theta_K \in \Re^K; \eta \in \Upsilon\right\}.$$

- Embedding: $H_0 \subset \mathcal{C}_K$.

- **Goodness-of-Fit Tests:** Score tests for

$$H_0 : \theta_K = 0, \beta \in \mathcal{B} \quad \text{versus} \quad H_1 : \theta_K \neq 0, \beta \in \mathcal{B}.$$

- Tests introduced in Peña (1998, JASA; 1998, AS).

- Since score tests, they possess local optimality properties.

- Likelihood function: obtained via product integration.

- Score function for θ_K at $\theta_K = 0$:

$$Q(\eta, \beta) = \sum_{i=1}^{n} \int_0^{\tau} \Psi_K(\eta) \left\{ \mathrm{d}N_i - Y_i \lambda_0(\eta) \exp\{\beta^{\mathrm{t}} X_i\} \mathrm{d}t \right\}.$$

- **Not a statistic** since η and β are unknown.

- Plug-in estimators for η and β under the restriction $\theta_K = 0$.

- Estimate β by $\hat{\beta}$ which solves

$$S(\beta) \equiv \sum_{i=1}^{n} \int_0^{\tau} [X_i - E(\beta)] \mathrm{d}N_i = 0;$$

$$E(t, \beta) = \frac{S^{(1)}(t, \beta)}{S^{(0)}(t, \beta)};$$

$$S^{(m)}(t, \beta) = \sum_{i=1}^{n} X_i^{\otimes m} Y_i \exp\{\beta^{\mathrm{t}} X_i\}, \quad m = 0, 1, 2.$$

- Estimate η by $\hat{\eta}$ which solves

$$R(\eta, \hat{\beta}) \equiv \sum_{i=1}^{n} \int_0^{\tau} \rho(\eta) \left\{ \mathrm{d}N_i - Y_i \lambda_0(\eta) \exp\{\hat{\beta}^{\mathrm{t}} X_i\} \mathrm{d}t \right\} = 0;$$

$$\rho(t, \eta) = \frac{\partial}{\partial \eta} \log \lambda_0(t, \eta).$$

- Test statistic:

$$S_K = \frac{1}{n}\left\{\hat{Q}^{\mathrm{t}}\right\}\left\{\hat{\Xi}^{-1}\right\}\left\{\hat{Q}\right\},$$

with

$$\hat{Q} = Q(\hat{\eta},\hat{\beta}) = \sum_{i=1}^{n}\int_0^{\tau}\Psi_K(\hat{\eta})\left\{\mathrm{d}N_i - Y_i\lambda_0(\hat{\eta})\exp\{\hat{\beta}^{\mathrm{t}}X_i\}\mathrm{d}t\right\}$$

- $\hat{\Xi}$ is an estimator of the limiting covariance matrix of $\frac{1}{\sqrt{n}}\hat{Q}$.

- S_K is a *function of the generalized residuals* $(R_1,\delta_1),\ldots,(R_n,\delta_n)$.

6. Asymptotics

- **Proposition:** If the parameters are known,

$$\frac{1}{\sqrt{n}}\begin{bmatrix} Q(\eta_0,\beta_0) \\ R(\eta_0,\beta_0) \\ S(\beta_0) \end{bmatrix} \xrightarrow{\mathrm{d}} N\left[\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} & 0 \\ \Sigma_{21} & \Sigma_{22} & 0 \\ 0 & 0 & \Sigma_{33} \end{pmatrix}\right],$$

so

$$\frac{1}{\sqrt{n}}Q(\eta_0,\beta_0) \xrightarrow{\mathrm{d}} N(0,\Sigma_{11}).$$

- **Theorem:** With estimated parameters,

$$\frac{1}{\sqrt{n}}\hat{Q} \equiv \frac{1}{\sqrt{n}}Q(\hat{\eta},\hat{\beta}) \xrightarrow{\mathrm{d}} N_K(0,\ \Xi)$$

where

$$\Xi = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} + (\Delta_1 - \Sigma_{12}\Sigma_{22}^{-1}\Delta_2)\Sigma_{33}^{-1}(\Delta_1 - \Sigma_{12}\Sigma_{22}^{-1}\Delta_2)^{\mathrm{t}}.$$

- **Proof:** Relies on Rebolledo's MCT.

7. Effects of the Plug-In Procedure

◇ From

$$\Xi = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} + (\Delta_1 - \Sigma_{12}\Sigma_{22}^{-1}\Delta_2)\Sigma_{33}^{-1}(\Delta_1 - \Sigma_{12}\Sigma_{22}^{-1}\Delta_2)^{\mathrm{t}},$$

replacing (η, β) by $(\hat{\eta}, \hat{\beta})$ to obtain the statistic $\hat{Q}$ has **no asymptotic effect** if

$$\Sigma_{12} = 0 \quad \text{and} \quad \Delta_1 = 0,$$

since Σ_{11} is the limiting covariance for $\frac{1}{\sqrt{n}}Q(\eta_0, \beta_0)$.

◇ Essence of "adaptiveness:" it does not matter that nuisance parameters (η, β) are unknown in $Q(\eta, \beta)$ since replacing them by their estimators does **not** make the asymptotic distribution of $Q(\hat{\eta}, \hat{\beta})$ different from $Q(\eta_0, \beta_0)$.

◇ $\Sigma_{12} = 0$ is an orthogonality condition between $\Psi_K(\eta_0)$ and $\rho(\eta_0)$.

◇ $\Delta_1 = 0$ is an orthogonality condition between $\rho(\eta_0)$ and $e(\eta_0, \beta_0)$.

◇ Orthogonality defined in a Hilbert space whose inner product is

$$\langle f, g \rangle = \int_0^\tau fg\nu_0(\mathrm{d}t),$$

with

$$\nu_0(A) = \int_A s^{(0)}(\eta_0, \beta_0)\lambda_0(\eta_0)\mathrm{d}t;$$

$$s^{(0)}(t; \eta_0, \beta_0) = \mathrm{plim}\frac{1}{n}S^{(0)}(t; \eta_0, \beta_0).$$

- ◇ Can we always choose Ψ_K to satisfy orthogonality conditions? **Yes,** via a Gram-Schmidt process ... but hard to implement!

- ◇ If orthogonality conditions are not satisfied, substituting $(\hat{\eta}, \hat{\beta})$ for (η_0, β_0) in $Q(\eta_0, \beta_0)$ impacts on the asymptotic distribution of $\hat{Q}$, *even if the estimators are consistent.*

- ◇ Second term in Ξ: effect of estimating η by $\hat{\eta}$.

- ◇ Third term in Ξ: effect of estimating β by the partial MLE $\hat{\beta}$.

- ◇ Estimating β by $\hat{\beta}$ leads to an increase in variance because this estimator is less efficient than the full MLE of β.

- ◇ Ignoring effect on variance could have **dire consequences** in the goodness-of-fit testing.

- ◇ If overall effect is a variance reduction, ignoring effect may result in a highly conservative test and lead into concluding model appropriateness when model is inappropriate.

8. Form of Test Procedure

- Recall:

$$\hat{Q} = Q(\hat{\eta}, \hat{\beta}) = \sum_{i=1}^{n} \int_0^{\tau} \Psi_K(\hat{\eta}) \left\{ \mathrm{d}N_i - Y_i \lambda_0(\hat{\eta}) \exp\{\hat{\beta}^{\mathrm{t}} X_i\} \mathrm{d}t \right\}.$$

- **Estimating Limiting Covariance Matrix, Ξ:** With

$$\hat{\Sigma} = \frac{1}{2n} \sum_{i=1}^{n} \int_0^{\tau} \begin{bmatrix} \Psi_K(\hat{\eta}) \\ \rho(\hat{\eta}) \\ X_i - E(\hat{\beta}) \end{bmatrix}^{\otimes 2} \left\{ \mathrm{d}N_i + Y_i \lambda_0(\hat{\eta}) \exp\{\hat{\beta}^{\mathrm{t}} X_i\} \mathrm{d}t \right\};$$

$$\hat{\Delta}_1 = \frac{1}{2n} \sum_{i=1}^{n} \int_0^{\tau} \Psi_K(\hat{\eta}) E(\hat{\beta})^{\mathrm{t}} \left\{ \mathrm{d}N_i + Y_i \lambda_0(\hat{\eta}) \exp\{\hat{\beta}^{\mathrm{t}} X_i\} \mathrm{d}t \right\};$$

$$\hat{\Delta}_2 = \frac{1}{2n} \sum_{i=1}^{n} \int_0^{\tau} \rho(\hat{\eta}) E(\hat{\beta})^{\mathrm{t}} \left\{ \mathrm{d}N_i + Y_i \lambda_0(\hat{\eta}) \exp\{\hat{\beta}^{\mathrm{t}} X_i\} \mathrm{d}t \right\};$$

an estimator of Ξ is

$$\hat{\Xi} = \hat{\Sigma}_{11} - \hat{\Sigma}_{12} \hat{\Sigma}_{22}^{-1} \hat{\Sigma}_{21} + (\hat{\Delta}_1 - \hat{\Sigma}_{12} \hat{\Sigma}_{22}^{-1} \hat{\Delta}_2) \hat{\Sigma}_{33}^{-1} (\hat{\Delta}_1 - \hat{\Sigma}_{12} \hat{\Sigma}_{22}^{-1} \hat{\Delta}_2)^{\mathrm{t}}.$$

- **Asymptotic α-Level Test:**

Reject $H_0 : \lambda(\cdot) \in \mathcal{C}; \beta \in \mathcal{B}$ whenever

$$S_K = \frac{1}{n} \left\{ \hat{Q}^{\mathrm{t}} \right\} \left\{ \hat{\Xi}^{-} \right\} \left\{ \hat{Q} \right\} \geq \chi^2_{K^*;\alpha},$$

$\hat{\Xi}^{-}$ = Moore-Penrose generalized inverse of $\hat{\Xi}$ and $K^* = \mathrm{rank}(\hat{\Xi})$.

9. Choices for Ψ_K

- Partition $[0, \tau]$: $0 = a_0 < a_1 < ... < a_{K-1} < a_K = \tau$

- Let

$$\Psi_K(t) = \left[I_{[0,a_1]}(t), I_{(a_1,a_2]}(t), ..., I_{(a_{K-1},\tau]}(t)\right]^{\mathrm{t}}.$$

- Then

$$\hat{Q} = [O_1 - E_1, O_2 - E_2, \ldots, O_K - E_K]^{\mathrm{t}}$$

  where

$$O_j = \sum_{i=1}^{n} \int_{a_{j-1}}^{a_j} \mathrm{d}N_i(t)$$

$$E_j = \sum_{i=1}^{n} \int_{a_{j-1}}^{a_j} Y_i(t) \exp\{\hat{\beta}^{\mathrm{t}} X_i\} \lambda_0(t; \hat{\eta}) \mathrm{d}t.$$

- O_j's are observed frequencies.

- E_j's are *estimated* dynamic expected frequencies.

- Extends Pearson's statistic, but 'counts are in a **dynamic** fashion.'

- Resulting test statistic **not** of form

$$\sum_{j=1}^{K} \frac{(O_j - E_j)^2}{E_j}$$

  because correction terms in variance destroys diagonal nature of covariance matrix.

- If adjustments are ignored, distribution is **not** chi-squared.

- Procedure extends Akritas (1988) and Hjort (1990).

- **Example:**

  - Consider no covariates (so $X_i = 0$);

  - $\mathcal{C} = \{\lambda_0(t;\eta) = \eta\}$, that is, Exponential distribution.

- For $j = 1, 2, \ldots, K$,

$$O_j = \sum_{i=1}^{n} \int_{a_{j-1}}^{a_j} \mathrm{d}N_i(t);$$

$$E_j = \hat{\eta} \sum_{i=1}^{n} \int_{a_{j-1}}^{a_j} Y_i(t)\mathrm{d}t;$$

$$\hat{\eta} = \frac{\sum_{i=1}^{n} \int_0^\tau \mathrm{d}N_i(t)}{\sum_{i=1}^{n} \int_0^\tau Y_i(t)\mathrm{d}t} = \frac{\text{Number of events}}{\text{Total exposure}}.$$

- Resulting test statistic:

$$S_K = E_\bullet \left[\mathbf{p} - \hat{\pi}\right]^{\mathrm{t}} \left[\mathrm{Dg}(\hat{\pi}) - \hat{\pi}\hat{\pi}^{\mathrm{t}}\right]^{-} \left[\mathbf{p} - \hat{\pi}\right],$$

  where

$$E_\bullet = \sum_{j=1}^{K} E_j;$$

$$\mathbf{p} = \frac{1}{E_\bullet}(O_1, O_2, \ldots, O_K)^{\mathrm{t}};$$

$$\hat{\pi} = \frac{1}{E_\bullet}(E_1, E_2, \ldots, E_K)^{\mathrm{t}}.$$

- $E_\bullet \neq n$.

- Appropriate degrees-of-freedom for the statistic is $K-1$ since $\mathbf{1}^{\mathrm{t}}\hat{\pi} = 1$.

- **Polynomial-type**:

$$\Psi_K(t;\eta) = \left[1,\ \Lambda_0(t;\eta),\ \Lambda_0(t;\eta)^2,\ \ldots,\ \Lambda_0(t;\eta)^{K-1}\right]^{\mathrm{t}}.$$

- Resulting $\hat{Q}$ vector has components $\hat{Q}_j, j = 1, 2, \ldots, K$ given by

$$\hat{Q}_j = \sum_{i=1}^{n} \exp\{-(j-1)\hat{\beta}X_i\} \int_0^{\tau} w^{j-1} \left\{\mathrm{d}N_i^R(w) - Y_i^R(w)\mathrm{d}w\right\},$$

where, with

$$R_i = \Lambda_0(Z_i;\hat{\eta}) \exp\{\hat{\beta}X_i\}$$

being the estimated residuals,

$$N_i^R(w) = I\{R_i \leq w; \delta_i = 1\};$$

$$Y_i^R(w) = I\{R_i \geq w\}.$$

- N_i^R's and Y_i^R's are generalized residual processes.

- Resulting test generalizes test proposed by Hyde's (mid '70's).

- **Other Choices:** Total-time-on-test ... this leads to a generalization of a test proposed by Barlow, Bartholomew, Brunk and Bremmer (1972) ... good for IFR alternatives.

10. Levels of Tests for Different K's

- **Polynomial:** $\Psi_K(t;\eta) = \left(1, \Lambda_0(t;\eta), \ldots, \Lambda_0(t;\eta)^{K-1}\right)^{\mathrm{t}}$

- $K \in \{1,2,3,4\}$; censoring proportion $\in \{25\%, 50\%\}$ and true failure time model were exponential and 2-Weibull. # of Reps = 2000

Null Dist.		Exponential(η)			
Parameters		$\eta = 2$		$\eta = 5$	
% Uncensored		75%	50%	75%	50%
Level		5%	5%	5%	5%
n	K				
50	2	4.65	6.65	5.00	5.60
	3	5.45	5.50	4.35	4.95
	4	6.55	5.50	5.10	5.85
	5	6.40	5.00	5.20	4.20
100	2	4.90	4.75	4.45	4.35
	3	4.55	4.35	4.65	4.25
	4	5.70	5.30	5.10	4.80
	5	5.75	4.90	5.30	4.75

Null Dist.		Weibull(α, η)			
Parameters		$(\alpha,\eta) = (2,1)$		$(\alpha,\eta) = (3,2)$	
% Uncensored		75%	50%	75%	50%
Level		5%	5%	5%	5%
n	K				
50	2	4.30	4.80	4.60	6.20
	3	5.40	5.15	6.30	5.70
	4	4.80	4.80	6.10	5.30
	5	5.25	3.45	6.70	4.60
100	2	3.90	4.95	4.20	4.80
	3	5.75	4.65	5.15	5.25
	4	5.55	4.15	5.00	4.30
	5	6.00	4.65	5.80	5.30

11. Power Function of Tests

Simulation Parameters: $n = 100$; 25% censoring.

Legend: Solid line ($K = 2$); Dots ($K = 3$); Short dashes ($K = 4$); Long dashes ($K = 5$).

Figure 1: Simulated Powers for Null: **Exponential** vs. Alt: **2-Weibull**

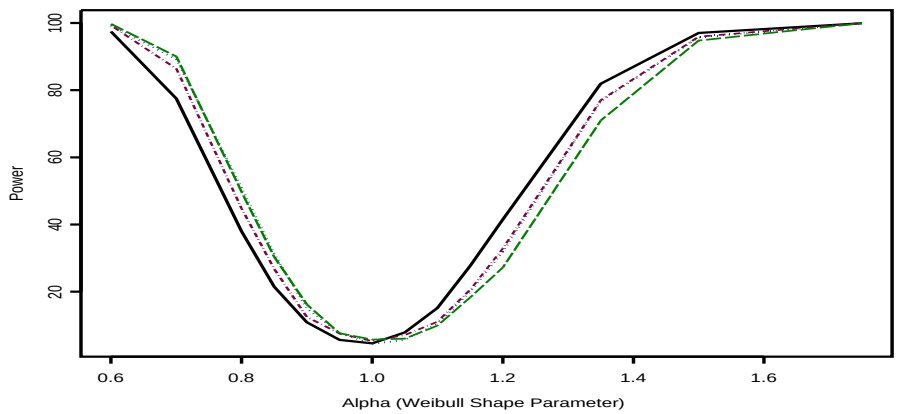


Figure 2: Simulated Powers for Null: **Exponential** vs. Alt: **2-Gamma**

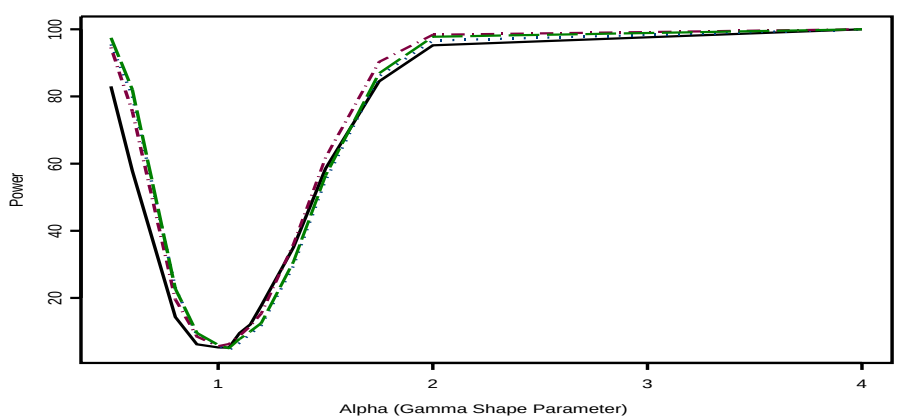
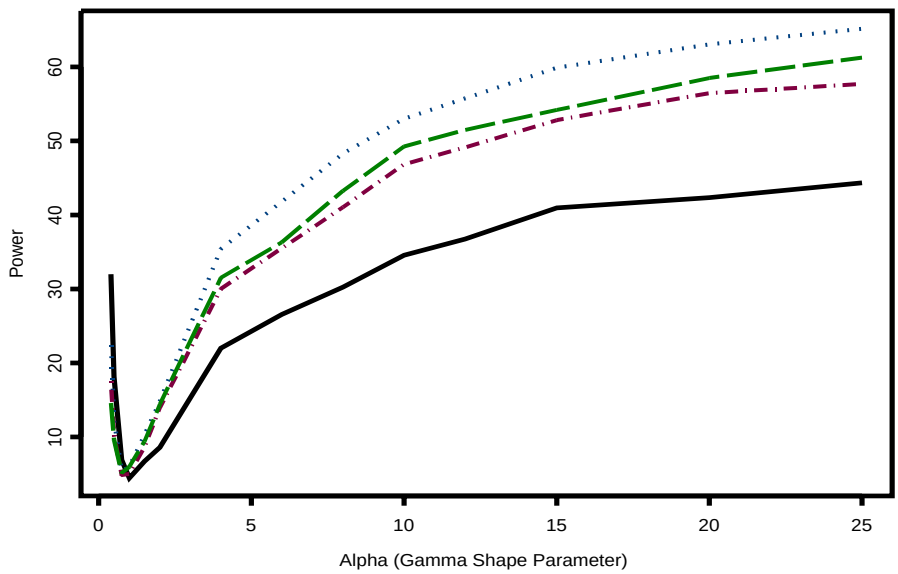


Figure 3: Simulated Powers for Null: **2-Weibull** vs. Alt: **2-Gamma**

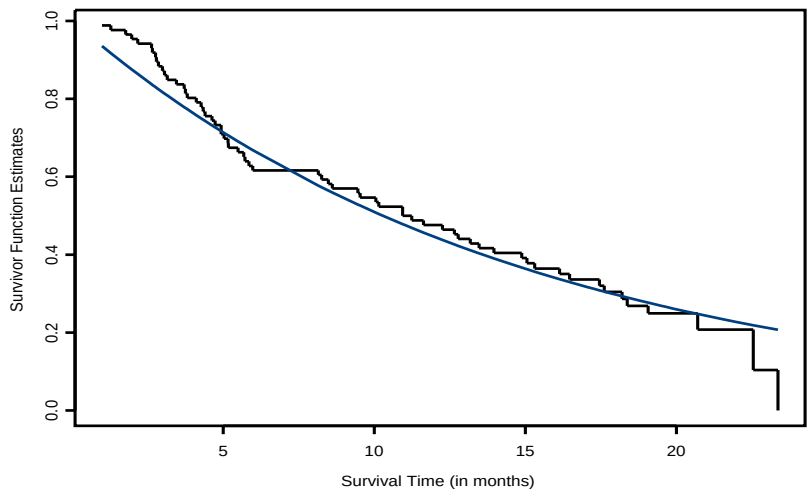


Some Observations:

- Appropriate smoothing order K depends on alternative considered.
- Not always necessary to have large K!
- Tests based on S_3 and S_4 could serve as omnibus tests, at least for the models in these simulations.
- Calls for a formal method to dynamically determine K. Plan: to utilize information-based criteria to determine appropriate smoothing order.

12. Back to the Lung Cancer Data

- Product-Limit Estimator (PLE) of survival curve together with confidence band, and the best fitting exponential survival curve.



- **Testing Exponentiality:** Values of test statistics using the polynomial-type specification, together with their p-values are:

$$S_2 = 1.92(p = .1661); \quad S_3 = 1.94(p = .3788);$$

$$S_4 = 7.56(p = .0561); \quad S_5 = 12.85(p = .0121).$$

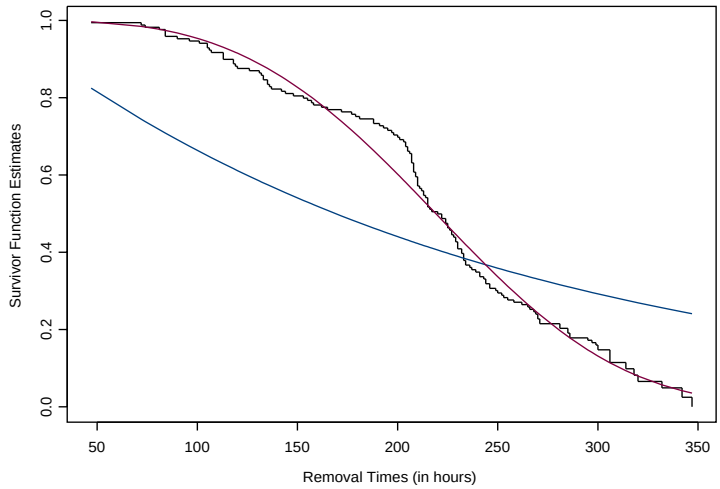
- Close to a constant function, but with high frequency terms.

- **Testing Two-Parameter Weibull:** Value of S_3, together with its p-value, is

$$S_3 = 8.35 \quad (p = .0153)$$

Values of S_4 and S_5 both indicate rejection of two-parameter Weibull model.

13. Back to the Car Tire Example



```
 Results for Testing Exponentiality

Summary of Results of Smooth Goodness-of-Fit Test
 Testing for Exponential Distribution

 Input FileName:  C:\Talks\TalkATUG\davislawrancecartiredata.txt
 Output FileName: C:\Talks\TalkATUG\cartire.out

 Sample Size =          171
 Sum of Z_i =    36457.0000000000
 Sum of Delta_i =          150
 Estimate of Eta =   4.114436185094769E-003
 Estimate of Mean =          243

 k      S_k            DF       P-Value
 1    0.0000           0        1.00000
 2  172.3710           1        0.00000
 3  174.9297           2        0.00000
 4  175.7617           3        0.00000
 5  175.7839           4        0.00000
```

Conclusion: Results very consistent with the graphical display: the exponential model does not hold.

```
 Results for Testing the Weibull Class

Summary of Results of Smooth Goodness-of-Fit Test
 Testing for the Two-Parameter Weibull Distribution
 Using the Polynomial Specification for Psi

 Input FileName:  C:\Talks\TalkATUG\davislawrancecartiredata.txt
 Output FileName: C:\Talks\TalkATUG\cartireWeibull.txt

 Sample Size =          171
 # of Uncensored Values =    150.000000000000
 Estimate of alpha =    3.41809555692639
 Estimate of eta =   4.085166006131305E-003

 k      S_k            DF       P-Value
 1    0.0000           0        1.00000
 2    0.5013           1        0.47891
 3    0.5550           2        0.75769
 4    6.4203           3        0.09286
 5    6.5183           4        0.16364
```

Conclusions From Analysis

- Cannot reject the hypothesis that the Weibull class holds for this car tire data.
- Notice in particular the p-values associated with the different smoothing order.
- Result of test consistent with the plots of the survivor estimates.

14. Extensions and Open Problems

- When failure times are discrete or interval=censored. Paper with V. Nair now in draft form. New twist is that embedding of the discrete hazards is through their odds!
- How about if the failure times are of a mixed-type, i.e., with continuous and discrete components?
- Problem of adaptively determining the smoothing order to be addressed.
- Promising possibility for $\mathbf{\Psi_K}$: **Wavelets as Basis??!**

15. Conclusions

- Presented a formal approach for developing GOF tests with censored data, and extends in a more *natural* way Neyman's smooth gof tests (cf., Rayner and Best (1989), *Smooth Goodness of Fit Tests*).
- Resulting tests possess optimality conditions by virtue of being score tests.
- Able to see effects of estimating nuisance parameters: (**Morale:** *Don't Ignore!*)
- Potential problems when using generalized residuals in model validation: (**Morale:** *Intuitive considerations may Fail!*)