STRUCTURAL INFERENCE REVISITED

an application to the generalized gamma model

Basílio de Bragança Pereira

Conferência de Estatística Indutiva

UFSCar – São Carlos SP

1-3 Setembro 2011

1- INTRODUCTION

Most process and experiments contain internal sources of variation, identifiable sources of variation that can be described by means of error variables (eq: error in the operation of a measuring instrument, the variation in the raw material to a process, variation due to randomization ingredient of an experimental desing).

These are the sources of variation in a system. Variation in a response variable derive from them.

Adequate description of a system neds a comprehensive model - a model with error variables to describe relevant aspects of sources of variation.

Classical model of statistics com be written

$$X$$
$$f(X,\emptyset)$$

Consist of a class o probability distributions indexed by a parameter $\emptyset$ and a response variable (random variable) X arising from a member of the class.

Structural model is written

$$X=\emptyset^{*}E$$
$$f(E)$$

Consists of a single probability distributions and a class of random variables index by $\emptyset$ and arising as an transformation on the error variable E. It is imposed in that the transformation (parameter) space is taken to be a group acting in the sample space.

The *error quantity* E is assumed to be a physical error and it can be described without refering to any particular parameter value Ø. Specifically the stronger assumption is made than the independence of E on Ø, in fact E itself is assumed unrelated to Ø.

For exemple, sucessive messurement erros, e, from a weighing scale are i.i.d., r.v. with pdf f(e) regardless of the item being weighet. Suppose the error is added to the weight Ø, of the item, to give the measured weight x=Ø+e and that n weighings (not necessarily distinct items) are to take place. If we view the erros e, ...$^{e}{}_{n}$ as stricly properties of the instrument that arise regardless of the order, shape, or weight of the items being weight, them the erros will be physical.

This assumption is close to the classical model

X

f(X-Ø)

Classical Model

- Does not describes any internal source of variation.

- Descrbes only external aspects of a system.

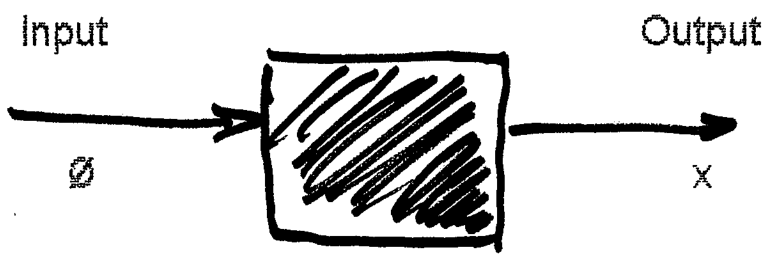
- Has an output or response variable, an input or parameter and a probability distribution describing the response variable for arbstrary values of the parameter.

- O casionally presents equations involving error variables ($Y=a+Bx+e$). Unfortunately does nat recognize or use a property of such equation hoving major consequences: from observations $Y_1 \ldots Y_n$ it is possible to calculate characteristics of the error values $e_1 \ldots e_n$.

- Need a variaty of reduction principles and optmality criteria. Only for special cases these methods reach unanimity for the solutions.

- Standard use of the classical model allows inference to be basead on observed x, on the model $f(x, \emptyset)$ and on nothing further.

It replaces the system by a <u>black box</u>. A black box is a model with imput variables, with autput variables, and with behavioral characteristics. But it is a model without linking between variables, without internal structure.



f(x: Ø)

Inferences based on classical model treat the observed x from process or system as an output from the black box with unknown input.

The black box has a direction of throughput: in by means of Ø, out by means of x.

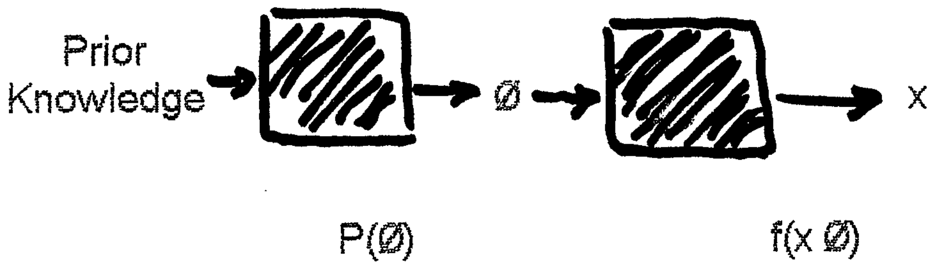
In practice, we have a response value and we are concerned with possible values of the input. This needs a black box with the oposite direction of throughput: in by means of x, out by means of Ø. The behavioral characteristic in one direction do not give characteristics for the oposite direction. For this, knowledge is needed concerning the internal structure of the system. This is not available in the classical model.

2- SOME REFINMENTS
Bayesian Method

- Withaws from the process and examines a compound - a classical model combined with probabilites taken as representing outside impressions concerning the unknowns.

- Introduces an ad ditional box that has output Ø with a frequency distribution p(Ø). The additional box is chosen to represent the investigator impression concerning the Ø value in the system.

- The x is an output from the combined boxes



Decision Theory

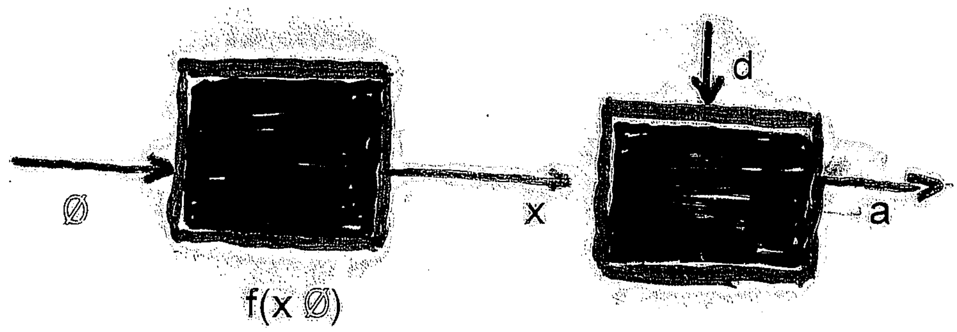
- Introduces an additional box. That has input x and an output a producing actions that might be appropriate to the context of the system.

- The additional box has a programing input d.

- A particular input d provides a frequency distribution for output a for each input x.

- The composite box is investigated to find programing input d that satisfy some optimality criterion relating inputs Ø and outputs a.

- The observed x from the system is used as an input to the additional box with its chosen programing input, the output a is the action concluding the investigation.



3- COMPREHENSIVE MODEL

- Is a stronger model, it describes additional characteristics of a process or experiment, the internal mechanism and structure.

- More substantial results than the classical model are obtained.

- The particular comprehensive model, The Structural Model: it is possible to calculate characteristics of the error variable, classical probability statements can be made.

- With he particular comprehensive model that is <u>Partly Structural</u> classical probabilitys statements can be made for some unknows, a marginal libelihood function can be derived for the remaining unknows.

- Information contained in the equations describing the structure of the system provides the base for making probability statements concerning the unknown inputs.

- The particular comprehensive model can be represented as white or transparent boxes:

Transparent Boxes
(Structural Models)

I) Simple Measurement Model

$\emptyset \longrightarrow \boxed{\emptyset + e} \longrightarrow X$

f(e)

II) Measurement Model

$(\emptyset\ V) \longrightarrow \boxed{\emptyset + Ve} \longrightarrow X$

f(e)

III) Measurement Model with Additional Quantity B

$(\emptyset, V) \longrightarrow \boxed{\emptyset + Ve} \longrightarrow X$

f(e,B)

The inferences on B are to be based on the marginal likelihood.

Pivotal Inference

(G. Barnard)

- Related to Fraser Structural Inference

- $x_1 \ldots\ldots x_m$

$$p_i = \frac{(x_i - \lambda)}{\sigma} \quad i = 1, 2 \ldots m \qquad \phi(\underset{\sim}{p}) = \left(1/\sqrt{2\pi}\right)^m \exp(-p'p/2)$$

$$\left.\begin{array}{ll} p_{m+1} = \lambda & \pi(\lambda) \\ p_{m+2} = \sigma & \pi(\sigma) \end{array}\right\} \text{Bayesian Pivotals.}$$

$$\phi(\underset{\sim}{p})\,\pi(p_{m+1})\,\pi(p_{m+2})$$

- Does not use or need group transformation

Barnard GA - 1985 - A coherent view of statistical inference. Univ. of Waterloo, Tech Report Series

STRUCTURAL INFERENCE FOR THE GENERALIZED GAMMA DISTRIBUTION

1. Introduction

The generalized gamma density for x > 0 can be written

$$f(x{:}\alpha,\beta,k)=\frac{\beta}{\Gamma(k)}\alpha^{-\beta k}x^{\beta k-1}\exp\left\{-\left(\frac{x}{\alpha}\right)^{\beta}\right\} \tag{1}$$

where $\alpha>0,\beta>0,k>0$. This model is often suggested as a lifetime distribution since it includes the widely used Exponential ($\beta=k=1$), Weibull ($k=1$), Gamma ($\beta=1$) and the Log-normal ($k\to\infty$). Other important distributions included are the Chi-Squared ($\alpha=2$, $\beta=1$, $k=n/2$), Chi $\left(\alpha=\sqrt{2},\beta=2,k=n/2\right)$, Half-Normal $\left(\alpha=\sqrt{2},\beta=2,k=1/2\right)$, Circular Normal $\left(\alpha=\sqrt{2},\beta=2,k=1\right)$, Spherical-Normal $\left(\alpha=\sqrt{2},\beta=2,k=3/2\right)$ and Rayleigh $\left(\alpha=C\sqrt{2},\beta=2,k=1 \text{ with } C>0\right)$.

It also has the properties that $Z=\lambda X$ has density $f(z{:}\lambda\alpha,\beta,k)$ and $Z=X^{m}$ has density $f(z{:}\alpha^{m},\beta/m,k)$ for $\lambda>0$ and $m\neq 0$.

2. Structural Model

The Structural Model (Fraser, 1968, ch 2) is described by two equations

$$X = \theta * W$$
$$f(\omega)d\omega$$

Consider now model (1). Making the transformation y=lnX and the reparametrization $\mu=\log\alpha$ and $\sigma=1/\beta$ we obtain the model

$$\begin{aligned} y &= \mu + \sigma\omega \\ f(\omega : k)d\omega &= \frac{1}{\Gamma(k)}\exp\left\{k\omega - e^{\omega}\right\} \end{aligned} \qquad (2)$$

A more general distribution is obtained by including the log normal and by letting $q = k^{-1/2}$ and allowing the error distribution at -q to be a reflexion about the origin of that at q, the general model becames (Farewell and Prentice, 1977)

$$y = \alpha + \sigma\omega$$

$$f(\omega : q)d\omega = \begin{cases} |q|\left(q^{-2}\right)^{q^{-2}} \exp\left\{q^{-2}\left(q\omega - \exp(q\omega)\right)\right\} / \Gamma\left(q^{-2}\right) & q \neq 0 \\ 2\pi^{-1/2}\exp\left(-\frac{1}{2}\omega^{2}\right) & q = 0 \end{cases} \qquad (3)$$

This is a location and scale model with a log-Gamma density for the error ω. So the model is in the form of the Conditional Structural Model with Additional Quantity q. (Fraser 1968, p. 188).

For q fixed and n observations $y' = (y_1,\ldots,y_n)$. The model is

$$y = \mu\, 1' + \sigma\omega \tag{4}$$

$$\prod_{i=1}^{n} f(\omega_i : q) \prod_{i=1}^{n} d\omega_i$$

where $1' = (1,\ldots,1)$ and $\omega' = (\omega_1,\ldots,\omega_n)$.

A suitable transformation variable or conditionally pivotal quantity is $\{\bar{y}, S_y\}$ where $\bar{y} = \sum_{i=1}^{n} y_i / n$ and $S_y = \left[\sum (y_i - \bar{y})^2 / (n-1)\right]^{1/2}$. The vector $d'(\omega) = d'(y) = (d_1,\ldots,d_n) = \left(\frac{y_1 - \bar{y}}{S_y}, \ldots, \frac{y_n - \bar{y}}{S_y}\right)$ is the corresponding orbital reference point or ancillary statistics.

If no assumption is made about the parameters (μ, σ) then observing $\{\bar{y}, S_y\}$ provides no information about the values of the error, ω, and quantities $\{\bar{\omega}, S_\omega\}$, in that whatever value of $\{\bar{\omega}, S_\omega\}$ is realized one can assert only $\{\bar{\omega}, S_\omega\} \in R x R^+$ on the basis of the observed y.

Inferences about $\{\overline{\omega}, S_\omega\}$ can be made for an assumed value for the quantitie q from the Reduced Structural Model (Fraser, 1968, p. 190).

$$\overline{y} = \mu + \overline{\omega} \qquad S_y = \sigma S_\omega$$

$$g(\overline{\omega}, S_\omega | \mathbf{d}: q) = K(\mathbf{d}: q) \prod_{i=1}^{n} f(\overline{\omega} + S_\omega d_i : q) S_\omega^n \frac{d\overline{\omega} dS_\omega}{S_\omega^2}$$

where K(**d**:q) is the integration constant. That is

$$g(\overline{\omega}, S_\omega | \mathbf{d}: q) = K(\mathbf{d}: q) \left(\frac{|q|}{\Gamma(q^{-2}) \left(\frac{1}{q^2} \right)^{\frac{1}{q^2}}} \right)^n \exp\left\{ \frac{1}{q^2} \left(qn\overline{\omega} - \sum_{i=1}^{n} e^{q(\overline{\omega} + S_\omega d_i)} \right) \right\} S_\omega^{n-2} d\overline{\omega} d \qquad (5)$$

The structural distribuition for (μ,σ) conditional in q is from (3) and (5) (Fraser, 1968, p. 190).

$$P(\mu,\sigma{:}\,q)=$$

$$=K(\mathbf{d}{:}\,q)\left(\frac{|q|}{\Gamma\left(q^{-2}\right)\left(\frac{1}{q^{2}}\right)^{\frac{1}{q^{2}}}}\right)^{n}\exp\left\{\frac{1}{q^{2}}\left(-nq\frac{\mu}{\sigma}+nq\frac{\bar{y}}{\sigma}-\sum_{i=1}^{n}e^{q\left(\frac{y_i-\mu}{\sigma}\right)}\right)\right\}\left(\frac{S_y}{\sigma}\right)^{n-1} \tag{6}$$

The marginal likelihood for q is

$$L(\mathbf{d}{:}\,q)\alpha\frac{1}{K(\mathbf{d}{:}\,q)} \tag{7}$$

where:

$$K(\mathbf{d}{:}\,q)=\int_0^{\infty}\int_{-\infty}^{\infty}\left(\frac{|q|}{\Gamma\left(q^{-2}\right)\left(\frac{1}{q^{2}}\right)^{\frac{1}{q^{2}}}}\right)^{n}\exp\left\{\frac{1}{q^{2}}\left(qn\bar{\omega}-\sum_{i=1}^{n}e^{q\bar{\omega}+S_{\omega}d_i}\right)\right\}S_{\omega}^{n-2}d\bar{\omega}dS_{\omega}$$

3. Inference

The structural distribution for (μ, σ) provides the basis for inferences concerning (μ, σ) for an assumed value o q. The marginal likelihood function provides the basis for inference concerning q.

3.1 Inference for the shape q.

Inference for the shape parameter q is obtained by ordering the values for q according to the relative marginal likelihood

$$R(d:q) = \frac{L(d:q)}{\sup_{q} L(d:q)} \tag{8}$$

The hypothesis $q = q_0$ can be tested from

$-2 \ln R(d:q_0)$

which has asymptotically a χ_1^2 distribution.

3.2. Tests of Significance For Location and Scale

Consider in general the measurement model (4). For the location parameter μ the hypothesis $\mu=\mu_0$ gives

$$\bar{y}=\mu_0+\sigma\bar{\omega},\ S_y=\sigma S_\omega \text{ and } t=\frac{\bar{\omega}}{S_\omega}=\frac{\sigma\bar{\omega}}{\sigma S_\omega}=\frac{\bar{y}-\mu_0}{S_y}$$

for the error characterist t.

This value of t is compared with the distribution of t derived from the error probability of $(\bar{\omega}, S_\omega)$ in (5). The jointly probability element for t and S_ω is

$$g(tS_\omega, S_\omega, d{:}q)\,dt\,dS_\omega$$

The marginal element for t is then (Fraser, 1968, p39):

$$g_L(t,d{:}q)dt=\int_0^\infty g(tS_\omega,S_\omega,d{:}q)dS_\omega dt=K(d{:}q)\int_0^\infty\prod_{i=1}^{n}f(tS_\omega+S_\omega d_i{:}q)S_\omega^{n-1}dS_\omega=$$

$$=K(d{:}q)\left(\frac{|q|}{\Gamma(q^{-2})\left(\frac{1}{q^2}\right)^{\frac{1}{q^2}}}\right)^{n}\int_0^\infty \exp\left\{\frac{1}{q^2}\left(qntS_\omega-\sum_{i=1}^{n}e^{q(tS_\omega+S_\omega d_i)}\right)\right\}S_\omega^{n-1}dS_\omega dt \qquad (9)$$

The hypothesis $\mu=\mu_0$ is assessed by comparing the calculated value t with the distributions of values described by $g_L(t,d{:}q)dt$ in (9).

For the scale parameter σ, the hypothesis $\sigma=\sigma_0$ gives $S_\omega = S_y/\sigma_0$ for the variable S_ω. This value for S_ω is compared with the distribution for S_ω obtained from the error probability distribution (Fraser, 1968; p39)

$$g_s(S_\omega, d:q) = K(d:q)\int_{-\infty}^{\infty}\prod_{i=1}^{n} f(\bar{\omega} + S_\omega d_i : q)\, d\bar{\omega}\, S_\omega^{n-2}\, dS_\omega =$$

$$= K(d:q)\left(\frac{|q|}{\Gamma(q^{-2})\left(\frac{1}{q^2}\right)^{\frac{1}{q^2}}}\right)^{n}\int_{-\infty}^{\infty}\exp\left\{\frac{1}{q^2}\left(qn\bar{\omega} - \sum_{i=1}^{n} e^{q(\bar{\omega}+S_\omega d_i)}\right)\right\} d\bar{\omega}\, S_\omega^{n-2}\, dS_\omega \tag{10}$$

The hypothesis $\sigma=\sigma_0$ is assessed by comparing the calculated value of S_y/σ_0 with the distributions of values described by $g_S(S_\omega, d:q)$ in (10).

4. Regression Model

Model (3) can be generalized to include explanatory variables $z=(z_1,\dots z_r)$, that is

$$y = z\beta + \sigma\omega$$
$$f(\omega:q) = |q|(q^{-2})^{q^{-2}} \exp\{q^{-2}(q\omega - \exp(q\omega))\} / \Gamma(q^{-2}) \tag{11}$$

which is in the form of a Conditional Regression Model with Additional Quantity q.

For q fixed and n observations $y'=(y_1,\dots,y_n)$ the model is

$$\mathbf{y} = \mathbf{Z}\beta + \sigma\boldsymbol{\omega}$$
$$\prod_{i=1}^{n} f(\omega_i:q)\prod_{i=1}^{n} d\omega_i \tag{12}$$

where $\omega = (\omega_1, \dots \omega_n)$, $\mathbf{Z}'$ is a r×n matrix defined by $\mathbf{Z}'=(z_1,\dots,z_n)$ and $\beta=(\beta_1,\dots,\beta_r)$ are regression coeficients.

A conditionally pivotal quantity is $(\hat{\beta}(y), S(y))$ where $\tilde{\beta}(y) = (\mathbf{Z'Z})^{-1}\mathbf{Z'Y}$ and $S^2(y) = |y - \mathbf{Z}\tilde{b}(y)|$ and the vector $\mathbf{D} = S^{-1}(y)(y - \mathbf{Z}\tilde{\beta}(y))$ is the corresponding ancillary statistics.

Inference about $\{\tilde{\beta}(\omega), S(\omega)\}$ for an assumed value of q can be obtained from the reduced model

$$\beta\,(y)=\beta+\sigma\beta(\omega)$$
$$S(y)=\sigma S(\omega) \tag{13}$$
$$g(\beta(\omega), S(\omega)|\mathbf{D}{:}q) = K(\mathbf{D}{:}q)\prod_{i=1}^{n} f\left(\sum_{u=1}^{r}\tilde{\beta}_u(\omega)z_{ui} + S(\omega)D_i\right)S^{n-r-1}\prod_{u=1}^{r} d\tilde{\beta}_u(\omega)\,dS(\omega)$$

where $K(\mathbf{D}{:}q)$ is the normalized constant.

From the error distribution (13) we can derive the structural distribution for (β,σ) for q fixed, the marginal likelihood for q and the location and scale distributions for significance tests (see Fraser, 1968, p. 126-131).

REFERENCES

Farewell,V.T. and Prentice,R.L.-1975-Interpreting the structural model. Statistische Hefte,v16,pp115-122.

Farewell,V.T. and Prentice,R.L.-1977-A study of distributional shape in life testing.Technometrics,v19,n1,pp69-75.

Fraser,D.A.S.-1968a-The Structure of Inference. Wiley.

Fraser,D.A.S.-1968b-A black box or a comprehensive model. Technometrics,v10,n2,pp219-229.