Credibility, hypothesis testing and regression software

Greg Taylor

Taylor Fry Consulting Actuaries
University of Melbourne
University of New South Wales

37th ASTIN Colloquium, 19-22 June 2007, Lake Buena Vista FL, USA

$\tilde{Y}$

Overview

- Simplest (Bühlmann) form of credibility formula

$$\tilde{Y} = (1-Z)\mu + ZY$$

where

Observable Y depends on latent parameter θ

$E[Y|\theta] = \mu(\theta)$

$E_\theta[\mu(\theta)] = \mu$

$\tilde{Y}$ estimates $\mu(\theta)$

Z = **credibility coefficient** = $\tau^2/(\tau^2 + \sigma^2)$

where

$\sigma^2 = E_\theta Var[Y|\theta]$

$\tau^2 = Var_\theta[\mu(\theta)]$

Question: how to estimate Z from data

$\tilde{Y}$

Bühlmann-Straub extension

- As before
  Observable Y depends on latent parameter θ
  $E[Y|\theta] = \mu(\theta)$
  $E_\theta[\mu(\theta)] = \mu$
  $\sigma^2 = E_\theta Var[Y|\theta]$
  $\tau^2 = Var_\theta[\mu(\theta)]$
- But now make observations $\{Y_{ij}, i=1,\ldots,I\}$ on each of J **risk classes** $j=1,\ldots,J$ characterised by $\theta=\theta_j$
- Credibility formula now estimates $\mu(\theta_j)$

$$\tilde{Y}_j = (1-Z)\bar{Y} + Z\,\bar{Y}_j$$

where Z is as before and

$$\bar{Y}_j = \Sigma_i Y_{ij} / I$$

$$\bar{Y} = \Sigma_{ij} Y_{ij} / IJ$$

- Data set is now $\{Y_{ij}, i=1,\ldots,I; j=1,\ldots,J\}$

Bühlmann-Straub extension

$Z = \tau^2/(\tau^2 + \sigma^2)$

where

$\sigma^2 = E_\theta Var[Y|\theta]$

$\tau^2 = Var_\theta[\mu(\theta)]$

Data set is $\{Yij, i=1,\ldots,I; j=1,\ldots,J\}$

- We might think of estimating

σ^2 by $\Sigma_j\{\Sigma_i [Y_{ij} - \bar{Y}_j]^2 /I\}/J$ (within-class variation)

and

$(\tau^2 + \sigma^2)$ by $\Sigma_{ij} [Y_{ij} - \bar{Y}]^2 /IJ$ (total variation)

Bühlmann-Straub extension

- Estimate σ^2 by within-class variation
- Estimate $(\tau^2 + \sigma^2)$ by total variation
- Hence estimate τ^2 by the difference, i.e. between-class variation
- We are performing an analysis of variance
- Note that

$$Z = \tau^2/(\tau^2 + \sigma^2) = 1/(1 + \sigma^2/\tau^2)$$

which is estimated by $1/(1 + 1/F)$ where $F =$

$$\frac{\text{between-class variation}}{\text{within-class variation}}$$

i.e. F-statistic of the analysis of variance (Zehnwirth, 1977)

Evaluation of the F-statistic

ANOVA model

$Y_{ij} = \alpha + \theta_j + \varepsilon_{ij}$, $i=1,\ldots,I$; $j=1,\ldots,J$

where the θ_j are parameters and the ε_{ij} centred random variables, iid for each fixed j

H_0: $\theta_j = 0$ for all j

H_1: $\theta_j \neq 0$ for at least one j

Calculate F-statistic using ANOVA or regression software

Regression formulation

$E[Y|\theta] = \mu(\theta) = \alpha + \theta$

Observations on $\theta = \theta_1, \theta_2, \ldots, \theta_J$

$Y_{ij} = \alpha + \theta_j + \varepsilon_{ij}$, $i=1,\ldots,I$; $j=1,\ldots,J$

$E[\varepsilon_{ij}] = 0$

Assume the ε_{ij} are iid

$\sigma^2 = E_\theta Var[Y|\theta] = Var[\varepsilon_{ij}]$

$E_\theta[\mu(\theta)] = E_\theta[\alpha+\theta] = b$

$\tau^2 = Var_\theta[\mu(\theta)] = Var_\theta[\theta]$

Regression formulation (cont'd)

$Y_{ij} = \alpha + \theta_j + \varepsilon_{ij}$, $i=1,\ldots,I; j=1,\ldots,J$

The ε_{ij} are iid with

$E[\varepsilon_{ij}] = 0$

$Var[\varepsilon_{ij}] = \sigma^2$

Write in matrix form

$$\underset{IJx1}{Y} = \underset{IJxJ}{X}\;\underset{Jx1}{\beta} + \underset{IJx1}{\varepsilon}$$

$E[\varepsilon] = 0$

$Var[\varepsilon] = \sigma^2 \underset{IJxIJ}{1}$

$$X = \begin{pmatrix} u_1 & & & \\ & u_2 & & \\ & & & \\ & & & u_J \end{pmatrix} \qquad \beta = \begin{pmatrix} \theta_1 \\ \theta_2 \\ \\ \theta_J \end{pmatrix}$$

where u_j is a vector consisting entirely of 1's

Regression formulation (cont'd)

Write in matrix form

$$\underset{IJ\times 1}{Y} = \underset{IJ\times J}{X} \ \underset{J\times 1}{\beta} + \underset{IJ\times 1}{\varepsilon}$$

$E[\varepsilon] = 0$

$Var[\varepsilon] = \sigma^2 \underset{IJ\times IJ}{1}$

$$X = \begin{pmatrix} u_1 & & & \\ & u_2 & & \\ & & & \\ & & & u_J \end{pmatrix} \qquad \beta = \begin{pmatrix} \theta_1 \\ \theta_2 \\ \\ \theta_J \end{pmatrix}$$

where u_j is a vector consisting entirely of 1's

Useful to re-parameterise to by choosing

$$X = \begin{pmatrix} u_1 & & & \\ u_1 & u_2 & & \\ & & & \\ u_1 & & & u_J \end{pmatrix} \qquad \beta = \begin{pmatrix} \theta_1 \\ \theta_2 - \theta_1 \\ \\ \theta_J - \theta_1 \end{pmatrix}$$

$= [X_0 \ \ X_+]$ where

$$X_0 = \begin{pmatrix} u_1 \\ u_1 \\ \\ u_1 \end{pmatrix} \qquad X_+ = \begin{pmatrix} 0 & & \\ & u_2 & \\ & & \\ & & u_J \end{pmatrix}$$

Regression formulation (cont'd)

Model form:

$$\underset{IJx1}{Y} = \underset{IJxJ}{X}\ \underset{Jx1}{\beta} + \underset{IJx1}{\varepsilon}$$

$$= [\underset{IJx1}{X_0}\ \ \underset{IJx(J-1)}{X_+}] \begin{pmatrix} \underset{1x1}{\beta_0} \\ \underset{(J-1)x1}{\beta_+} \end{pmatrix} + \underset{IJx1}{\varepsilon}$$

$$= X_0\,\beta_0 + X_+\,\beta_+ + \varepsilon$$

H_0: $\beta_+=0$

H_1: $\beta_+\neq 0$ [Bühlmann-Straub]

Calculate F-statistic for this regression test

Then can prove that $Z = 1/(1 + 1/F)$ is Bühlmann-Straub credibility coefficient

Other models

- We have obtained a simple procedure for estimating the credibility coefficient in a Bühlmann-Straub type of model
- But this is a very simple model
- Do the same ideas extend to other models? e.g.
  - Hachemeister regression model
  - Taylor hierarchical model

General framework

General framework will be Hachemeister's (paper is slightly more general)

$$\underset{nx1}{Y} = \underset{nxp}{X}\ \underset{px1}{\beta} + \underset{nx1}{\varepsilon} \qquad E[\varepsilon] = 0,\ Var[\varepsilon] = V$$

$E[\beta] = b$, $Var[\beta] = \Gamma$

Define $\hat{\beta}$ as the standard general linear regression (fixed effects) estimate of β:

$$\hat{\beta} = (X^T V^{-1} X)^{-1} X^T V^{-1} Y$$

Hachemeister credibility estimate of $E[X\beta|Y]$ is:

$$\tilde{Y} = (1-Z)\, Xb + Z\, X\hat{\beta}$$

where

$$Z = [1 + (X\Gamma X^T V^{-1})^{-1}]^{-1}$$

General framework (cont'd)

$$\underset{n\times 1}{Y} = \underset{n\times p}{X}\ \underset{p\times 1}{\beta} + \underset{n\times 1}{\varepsilon} \qquad E[\varepsilon] = 0,\ Var[\varepsilon] = V$$

$$E[\beta] = b,\ Var[\beta] = \Gamma$$

$$X = [\underset{n\times p_0}{X_0}\ \ \underset{n\times(p-p_0)}{X_+}]$$

$$\underset{p\times p}{\Gamma} = \begin{pmatrix} \underset{p_0\times p_0}{\Gamma_0} & * \\ * & \underset{(p-p_0)\times(p-p_0)}{\Gamma_+} \end{pmatrix}$$

$V = \sigma^2 W \qquad \Gamma = \tau^2 G$

where W, G are known and σ^2, τ^2 are unknown

Then

$$Z = [1 + (X\Gamma X^T V^{-1})^{-1}]^{-1}$$

$$= [1 + (v\ XGX^T W^{-1})^{-1}]^{-1}$$

where $v = \tau^2/\sigma^2$

Alternative regression hypotheses

$Y = X \ \beta + \varepsilon$

nx1 nxp px1 nx1

$X = [X_0 \quad X_+]$

nxp_0 $nx(p\text{-}p_0)$

Hypothesis H_1: $Y = X\beta + \varepsilon$, write $Y = X_1\beta_1 + \varepsilon$

Hypothesis H_0: $Y = X_0\beta_0 + \varepsilon$

Define

$\hat{Y}_i = X\,\hat{\beta}_i$ [fitted values]

$RSS = (Y - \hat{Y}_1)^T W^{-1}(Y - \hat{Y}_1)$ [residual sum of squares]

$SS_{reg} = (\hat{Y}_1 - \hat{Y}_0)\ W^{-1}(\hat{Y}_1 - \hat{Y}_0)$ [regression sum of squares]

$F = [SS_{reg}/(p-p_0)] / [RSS/(n-p)]$ [regression F-statistic]

Estimation of credibility parameter

$Z = [1 + (v\ XGX^T W^{-1})^{-1}]^{-1}$

where $v = \tau^2/\sigma^2$

Estimate v by

$\hat{v} = \max[0, K(F-1)]$

for constant K:

$K = (p - p_0) / \mathrm{tr}\ [X_+^T W^{-1} (1 - P_0) X_+ G]$

with $P_0 = X_0\ (X_0^T W^{-1} X_0)^{-1}\ X_0^T W^{-1}$

Numerical evaluation of credibility parameter

- Formulate credibility problem in regression terms
  - Some (or all) of the regression parameters will be subject to priors
- Identify null hypothesis H_0 as that in which all the parameters subject to priors are set to zero
- Calculate F-statistic for testing the regression model (H_1) against H_0
  - All steps up to this point may be carried out using regression software
- Estimate credibility parameter v on the basis of this F
- Hence compute credibility matrix Z

Examples

- The structure investigated here covers many of the credibility models of interest
  - Hachemeister regression model
  - Bühlmann-Straub model
  - Taylor hierarchical credibility model
  - Sundt hierarchical regression model

Hachemeister regression example

$X = [X_0 \quad X_+]$

$$X_0 = \begin{pmatrix} u_m & t_m \\ u_m & t_m \\ \vdots & \vdots \\ u_m & t_m \end{pmatrix} \qquad X_+ = \begin{pmatrix} 0 & 0 & \dots & 0 & 0 \\ u_m & t_m & \ddots & & \\ \vdots & \vdots & & \ddots & \\ 0 & 0 & \dots & u_m & t_m \end{pmatrix}$$

where each pair u_m, t_m represents a state with

$\underset{m\text{x}1}{u_m} = [1,1,\dots,1]^T$ and $\underset{m\text{x}1}{t_m} = [1,2,\dots,m]^T$ is a trend factor

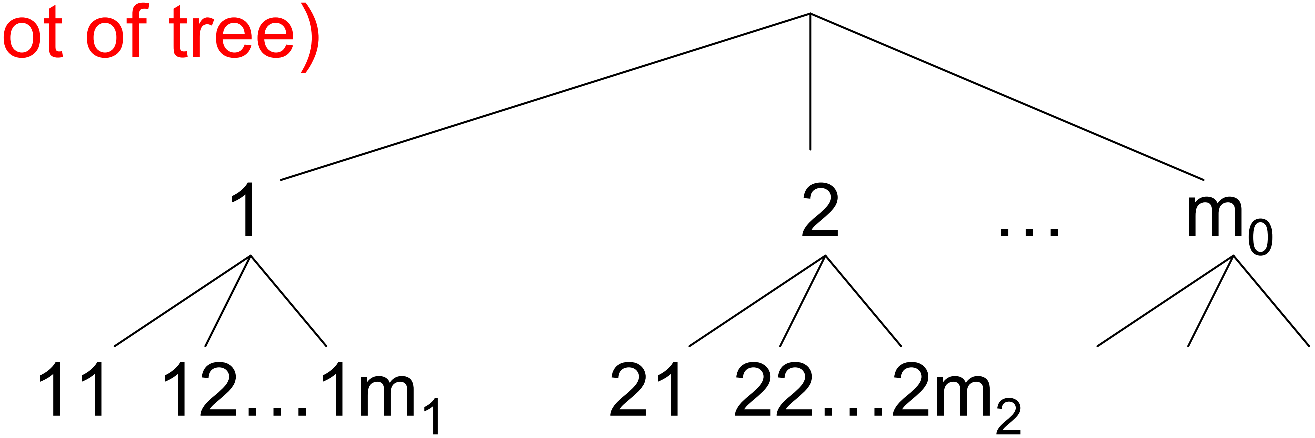
Sundt hierarchical model

- Risk classes form a tree structure with q levels
- Nodes at level k denoted $j_1 j_2 \ldots j_k$

Level 0 (root of tree)

Level 1

Level 2



Level 3 etc

Sundt hierarchical model (cont'd)

- Hachemeister regression structure at each node of the tree
- Consider any one path through the tree from root to leaf
- Denote its nodes 1,2,…,q
- Take a vector observation Y_j at node j
- Suppose there is a latent parameter θ_j at node j
- Assume
  - Y_j and θ_k are conditionally independent given $\theta_1,\ldots,\theta_j$ if k>j
  - $[Y_1^T,\ldots,Y_j^T]^T$ and $[Y_{j+1}^T,\ldots,Y_q^T]^T$ are conditionally independent given $\theta_1,\ldots,\theta_j$, j=1,…,q-1
  - $E[Y_j|\,\theta_1,\ldots,\theta_j] = X_j\,\beta_j(\theta_1,\ldots,\theta_j)$, j=1,2,…,q for non-stochastic matrix X_j and vector-valued function $\beta_j(.)$
  - $E[\beta_j(\theta_1,\ldots,\theta_j)|\,\theta_1,\ldots,\theta_{j-1}] = \beta_{j-1}(\theta_1,\ldots,\theta_{j-1})$, j=1,2,…,q and $E[\beta_1(\theta_1)] = \beta_0$

Sundt hierarchical model (cont'd)

Credibility estimator of $E[\beta_j(\theta_1,\ldots,\theta_j)|\theta_1,\ldots,\theta_{j-1}]$ is

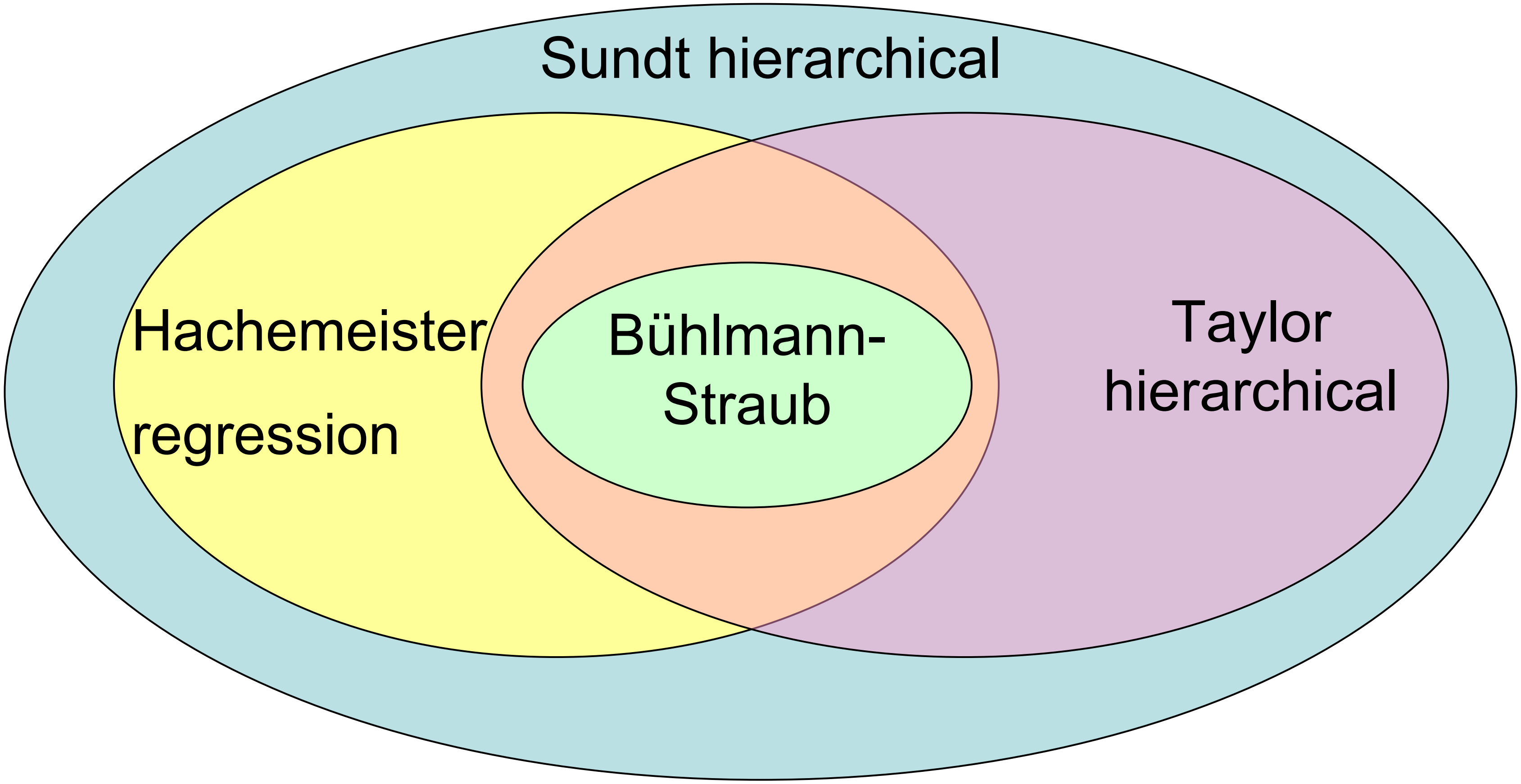
$$\tilde{\beta}_j = (1-Z_j)\,\tilde{\beta}_{j-1} + Z_j\,\hat{\beta}_j,\ j=1,2,\ldots q$$

$$\tilde{\beta}_0 = \beta_0$$

where $\hat{\beta}_j$ is the conventional generalised least squares estimator of $E[\beta_j(\theta_1,\ldots,\theta_j)|\theta_1,\ldots,\theta_{j-1}]$ based on data $Y_j,\ldots,Y_q$ and Z_j has the regression credibility matrix form as earlier

Z_j is estimated by same procedure as earlier

Relation between models



Conclusion

- Formulation of a credibility problem in regression terms and use of the F-statistic for testing the signifiance of the regression reduces the computation of the credibility coefficient (matrix) to a routine
- This formulation can be applied to a wide variety of credibility models
- It enables the use of regression software for ease of computation of the credibility coefficient