重回帰モデルによる安全管理のシミュレーション

熊本大学工学部　正　員　三池　亮次
熊本大学工学部　学生員　○市来　昌彦
熊本大学工学部　正　員　小林　一郎
熊本大学工学部　学生員　山城　久富

1．まえがき．　アーチダムの竣工後，長期間が経過し，ダムの挙動が安定しており可逆的な変形を保ちつつあるとき，例えば，大地震があったとする。地震前後のアーチダムの挙動に異常が発生したかどうかを検出する一手法として，さきに重回帰における差の検定理論に基づく解析法を提案した[1]この理論は重回帰における回帰偏差が正規分布に従うことが前提であった。回帰偏差が正規分布に従わなくてもデータサイズが十分に大きい場合には，中心極限定理に従って回帰推定値は正規分布に収れんし，差の検定理論は成立する。回帰推定値が中心極限定理に従って正規分布に収れんするのに必要なデータサイズを推定するために，我々はまた，コンピュター・シミュレーション解析の一手法であるブートストラップ法をこの問題に適用した[2]その結果，回帰偏差が指数分布のような非対称分布する場合でも回帰推定値は正規分布に近づくことが確認された。ただし，ブートストラップ法では，重回帰モデルの説明変数が多い場合，中心極限定理が成立するためには，データサイズは，かなり大きいことが必要であることしか分からなかった。ブートストラップ手法自身に若干の疑義のあることも指摘された。

本講では，次のようなコンピュター・シミュレーションを通して，比較的にデータサイズが小さく回帰偏差が正規分布に従わなくも，重回帰における差の検定理論に基づく安全管理の手法が有効であることを明らかにする。すなわち最初に重回帰モデルの説明変数$\mathbb{X}$，回帰係数$\mathbb{B}$と指数または正規分布に従う回帰偏差$\mathbb{e}$の予備と管理区間の２組の値を与え，重回帰式の応答量$\mathbb{Y}$の各組の値を求める。次に回帰係数と回帰偏差の値は未知として逆に説明変数$\mathbb{X}$と応答変量$\mathbb{Y}$より回帰係数の推定値$\hat{\mathbb{B}}$を求める。また$\mathbb{Y}$の推定値$\hat{\eta}=\mathbb{X}\hat{\mathbb{B}}$を求める。従って，応答変量$\mathbb{Y}$の真値$\eta=\mathbb{X}\mathbb{B}$が予備と管理区間で差があるかどうかを検定する問題である。

2．差の検定理論　予備データ（$l=1$）と管理データ（$l=2$）に対して重回帰モデルが

$$\mathbb{Y}_1=\mathbb{X}_1\mathbb{B}_1+\mathbb{e}_1$$

であるとする。$\mathbb{B}_1$の最小二乗推定値は

$$\hat{\mathbb{B}}_1=\mathbb{S}_1^{-1}\mathbb{X}_1^{T}\mathbb{Q}_1\mathbb{Y}_1\quad,\qquad \mathbb{S}_1=\mathbb{X}_1^{T}\mathbb{Q}_1\mathbb{X}_1$$

残差$\mathbf{v}_1$と残差平方和RSS_1は，

$$\mathbf{v}_1=\mathbb{Y}_1-\hat{\eta}_1\quad,\ RSS_1=\mathbb{Y}_1^{T}\mathbb{Q}_1\mathbb{Y}_1-\hat{\mathbb{B}}_1^{T}\mathbb{X}_1^{T}\mathbb{Q}_1\mathbb{Y}_1$$

回帰偏差と応答変量の推定値$\hat{\eta}_1$のｊ番目の値の不偏分散$\hat{\sigma}^2_{1,j}$は

$$\hat{\sigma}_1^2=\frac{RSS_1}{n_1-p'}\quad,\ p'=p+1\quad,\ \hat{\sigma}^2_{\eta 1,j}=D_{1,j}\hat{\sigma}_1^2$$

となる。ここに，p'は，定数項を含む説明変数の数，$D_{1,j}$は$\mathbb{X}_1\mathbb{S}_1^{-1}\mathbb{X}_1^{T}$のｊ番目の対角要素である。

管理区間の説明変数$\mathbb{X}_2$に$\mathbb{B}_1$を乗じた応答変量$\eta_2'=\mathbb{X}_2\mathbb{B}_1$と$\mathbb{B}_2$を乗じた$\eta_2=\mathbb{X}_2\mathbb{B}_2$に差があるかどうかを検定する。もし偏差$\mathbb{e}_1$が正規分布$N(\mathbb{O},\sigma_1^2\mathbb{I})$に従うなら上記の各応答量のｊ番目の推定値$\hat{\eta}_{2,j}=\mathbf{x}_{2,j}^{T}\mathbb{B}_2\sim N(\eta_{2,j},D_{2,j}\sigma_2^2)$，$\hat{\eta}_{2,j}'=\mathbf{x}_{2,j}^{T}\hat{\mathbb{B}}_1\sim N(\eta_{2,j}',D_{2,j}'\sigma_1^2)$である。ここに，$\mathbf{x}_{2,j}^{T}$は$\mathbb{X}_2$のｊ番目の行ベクトルで$D_{2,j}=\mathbf{x}_{2,j}^{T}\mathbb{S}_2^{-1}\mathbf{x}_{2,j}$，$D_{2,j}'=\mathbf{x}_{2,j}^{T}\mathbb{S}_1^{-1}\mathbf{x}_{2,j}$である。$\eta_{2,j}$，$\eta_{2,j}'$の間に差がない，すなわち，予備と管理データに対して$\mathbb{B}_1=\mathbb{B}_2$，$\sigma_1^2=\sigma_2^2$である仮設の下で，

$$T\leq t\,(n_1-p'\,;0.05)$$

ここに，$t\,(n_1-p'\,;0.05)$は自由度n_1-p'のｔ分布の両側信頼係数５%の値であり，Tは，

$$T=\frac{\hat{\eta}_{2,j}'-\hat{\eta}_{2,j}}{(D_{2,j}'+D_{2,j})\,\hat{\sigma}_1^2}$$

である。かくして上式において，$t\,(n_1-p'\,;0.05)$を管理区間限界とするＴの管理図を描くことができ，これによって我々は，予備と管理区間で差があるかどうかを検定が可能となる。

3．重回帰モデルによる差の検定シミュレーション解析

以下のような，重回帰モデルによる差の検定安全管理のシミュレーション解析を行った。

STEP 1　説明変数p'を３，５，７個の３通り，サンプルサイズｎを5,10,20,40,60 個の５通りの重回帰モ

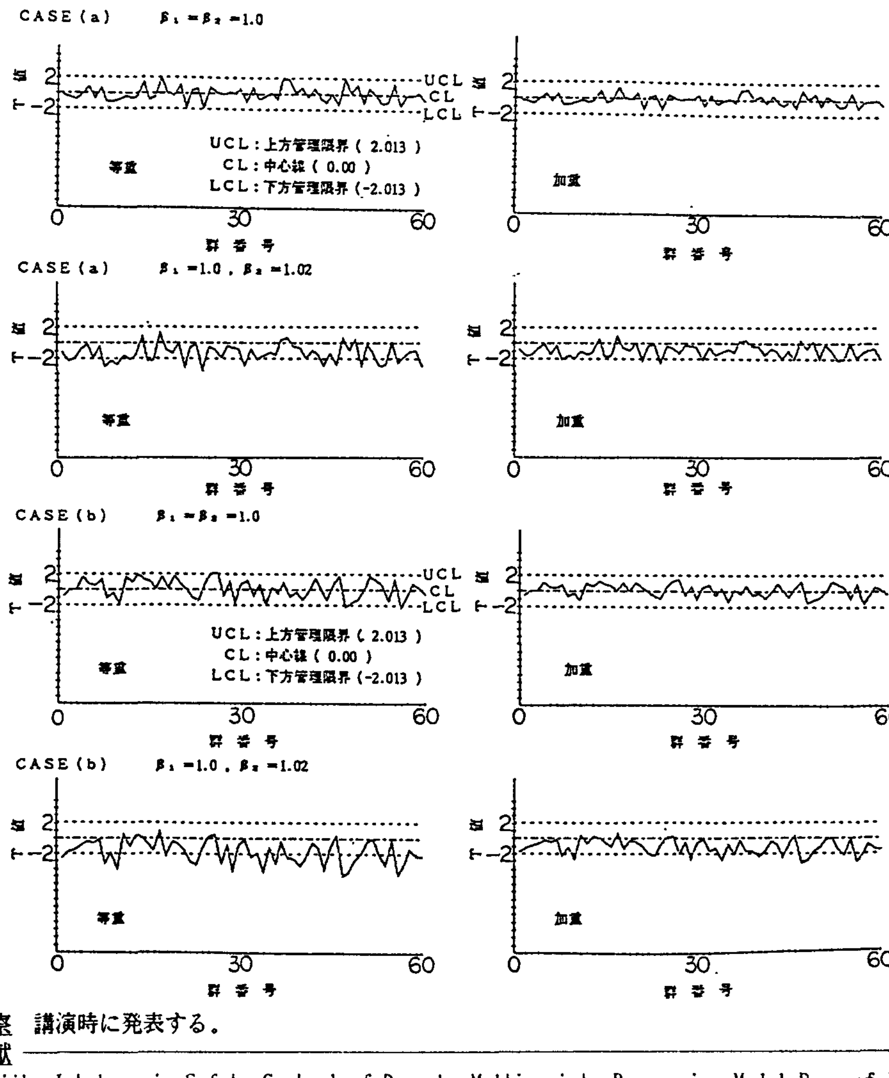
デルを設定する。その予備区間における回帰係数$\beta_{1,i}=1.0$ $(i=0,1,2\cdots)$管理区間の回帰係数$\beta_{2,i}=1.0, 1.02$の２通りで既知であるとする。

STEP 2　予備と管理区間の回帰偏差の確率分布が次のCASEの何れかであるとして，その値を与える。上記の何れかの組み合わせに対し重回帰モデルを設定。次に，回帰係数と回帰偏差の確率分布は未知として，回帰分析を行い，予備と管理区間における差が正しく検定されるかどうかのシミュレーション解析をおこなう。

ＣＡＳＥ（ａ）：予備と管理区間の各偏差$\mathbf{e}_1$，$\mathbf{e}_2$は，それぞれ，母平均 0，母分散1.0 の正規分布$N〔0,1〕$，および$N〔0,4.0〕$に従う。

ＣＡＳＥ（ｂ）：偏差$\mathbf{e}_1$ $\mathbf{e}_2$は，$\mathbf{e}_i=\alpha\log_e(1-u_i)-\alpha\log_e 2$の指数分布に従う。ただし，$\mathbf{e}_1$では，$\alpha=1.0$ $(\sigma_1^2=1.0)$，$\mathbf{e}_2$では，$\alpha=2.0$ $(\sigma_1^2=4.0)$とする。

予備と管理の両区間にそれぞれ，$\Omega_i=\mathbb{I}$ $(i=1,2)$の等重の場合と予備と管理区間における分散の相違による影響を除去するために両区間に対して重さΩ_iが異なるとして加重差の検定の場合についてのシミュレーション解析の結果は次の図のとおりである。ただし，$p'=7$，$n=60$の場合の解析結果を示す。



4．考察　講演時に発表する。

参考文献

1）R.Miike, I.kobayasi：Safety Control of Dams by Multivariate Regression Model, Proc.of ICOSSAR. 1985.　2）三池，小林：重回帰安全管理へのブートストラップ手法の適用，構造物の安全性および信頼性，Vol.1.1，第一回構造物の安全性と信頼性に関する国内シンポジューム，1987.