The chain ladder and Tweedie distributed claims data

Greg Taylor

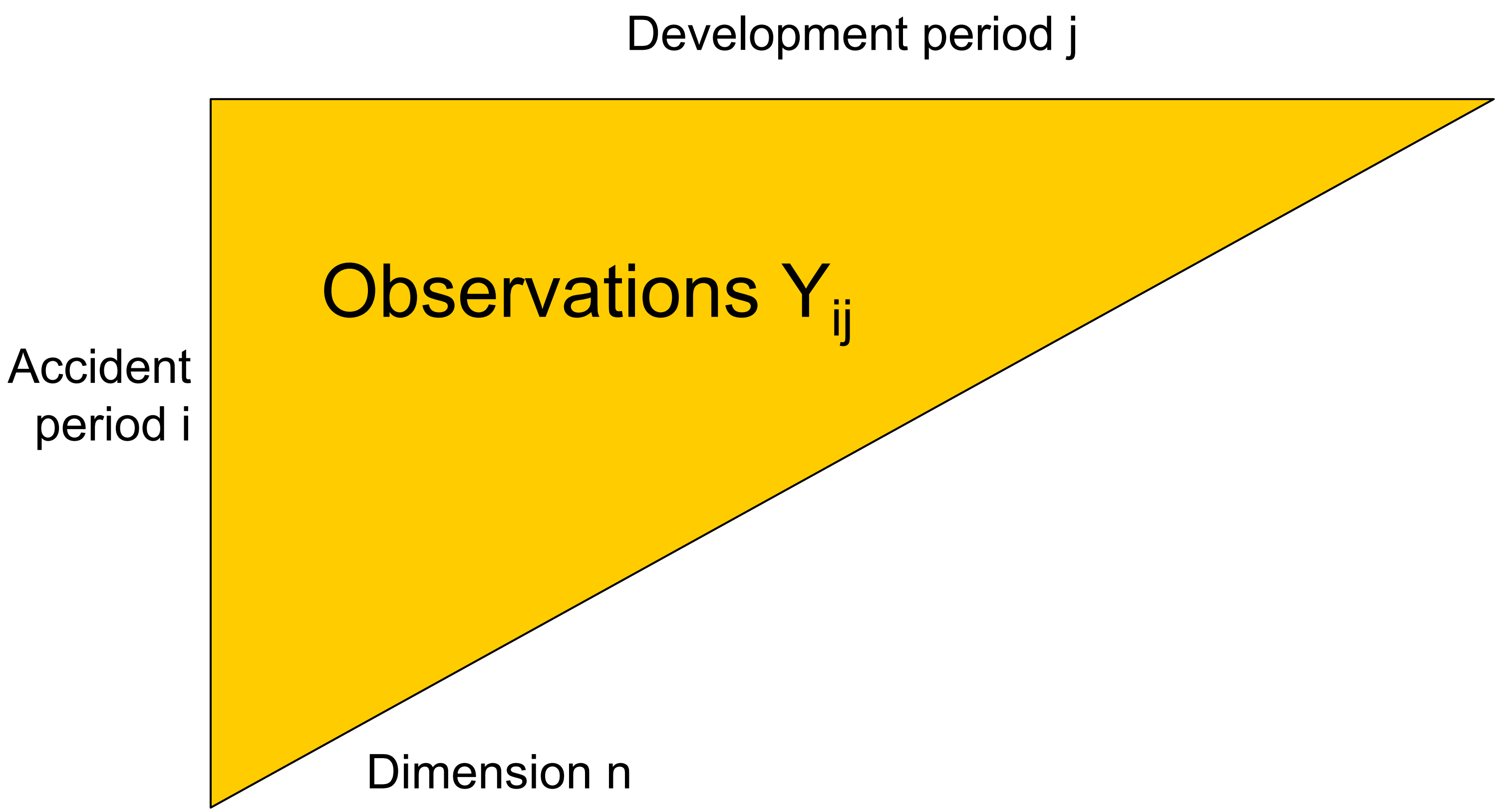
Taylor Fry Consulting Actuaries
University of Melbourne
University of New South Wales

Casualty Actuarial Society
Spring Meeting 3-6 May 2009, New Orleans

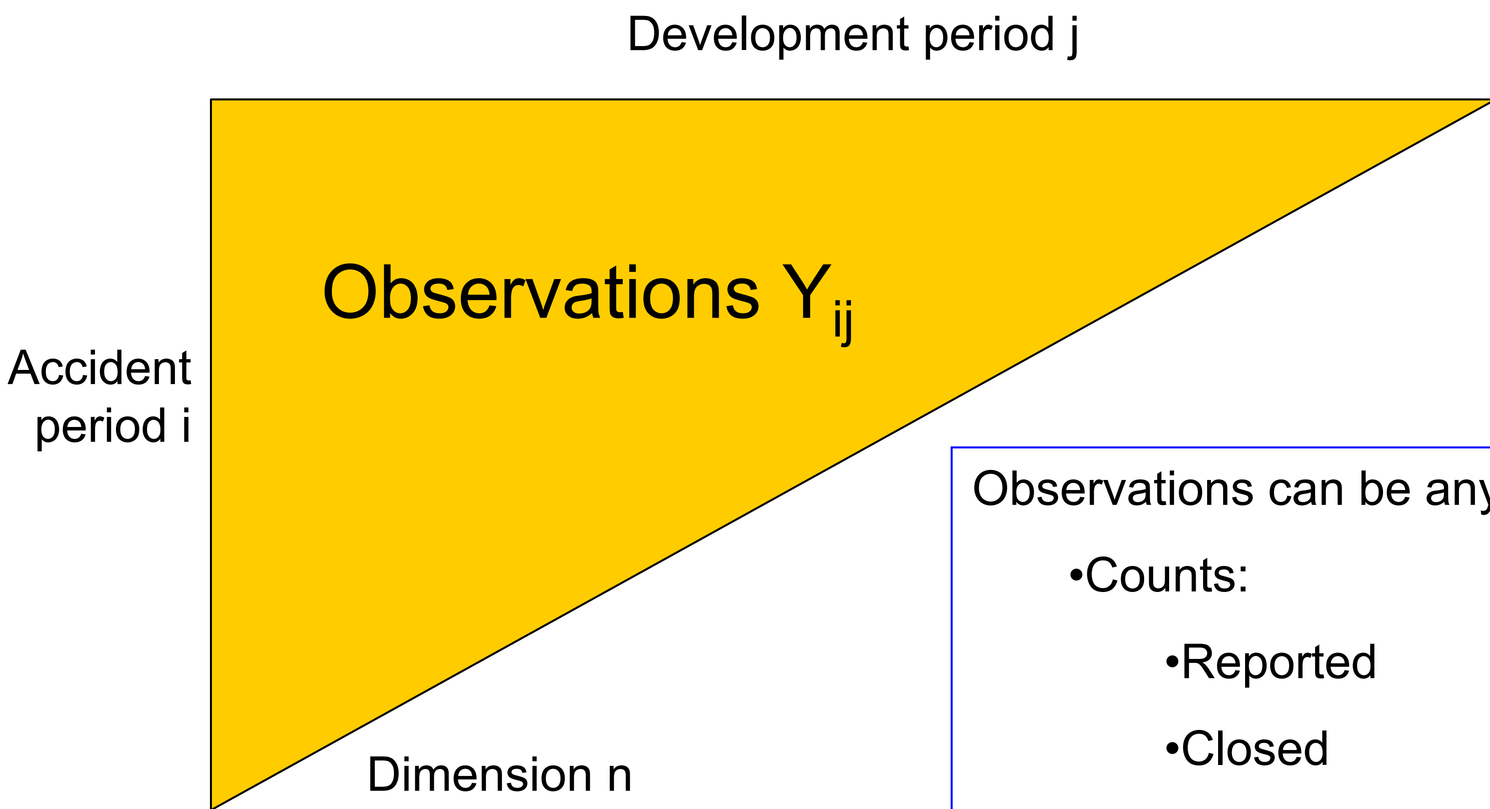
Purpose and overview

- Chain ladder
  – When is it maximum likelihood and when not?
  – When it isn't, is it close to ML?
- Questions considered in the context of the Tweedie family of distributions for chain ladder observations

Framework


Development period j
Observations Y_{ij}
Accident period i
Dimension n

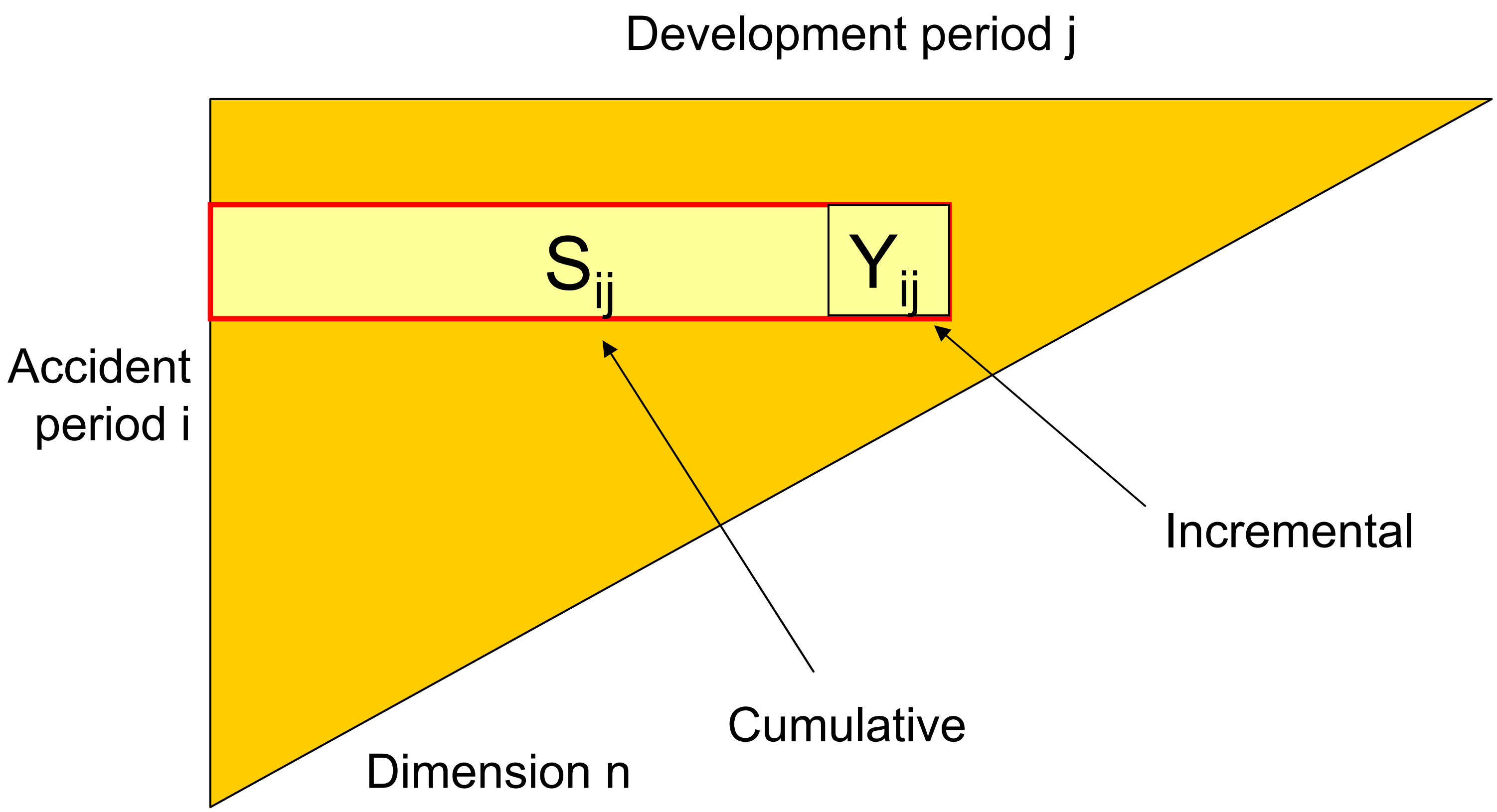
Framework



Observations can be anything:

- Counts:
  - Reported
  - Closed
- Amounts:
  - Paid
  - Incurred

Framework



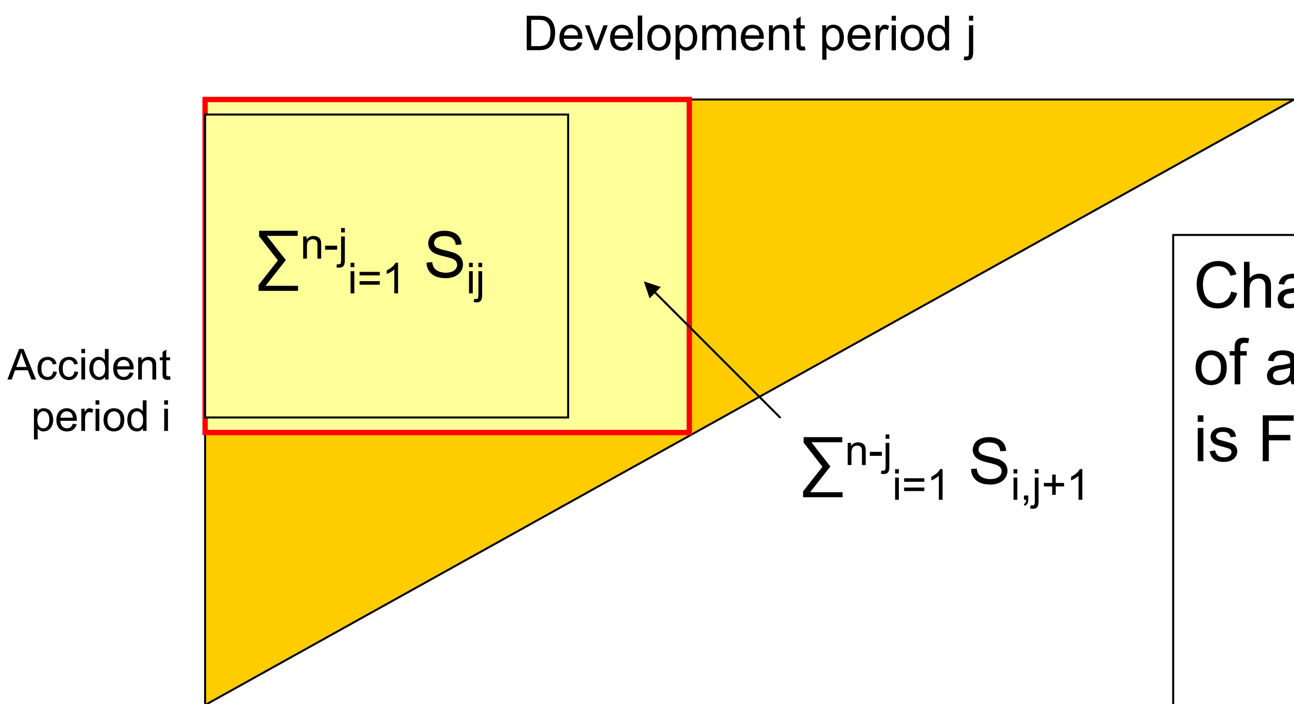
Development period j
S_{ij}
Y_{ij}
Accident period i
Incremental
Cumulative
Dimension n

Chain ladder model

- Model formulation (due to Mack (1991, 1993)):
  - **Assumption CL1:** $\mathbf{E[S_{i,j+1} \mid S_{i1},S_{i2}\ldots,S_{ij}] = S_{ij}f_j}$, independently of i for some set of parameters f_j (age-to-age factors)
  - **Assumption CL2:** Rows of the data triangle are stochastically independent, i.e. Y_{ij} and Y_{kl} are independent for $i \neq k$
- **NOTE:** chain ladder is **distribution free**
  - No assumption about distribution of Y_{ij}

Taylor Fry
Consulting Actuaries

Chain ladder estimation



Chain ladder estimate of age-to-age factor f_j is F_j =

$$F_j = \frac{\sum^{n-j}_{i=1} S_{i,j+1}}{\sum^{n-j}_{i=1} S_{ij}}$$

A slightly different model

- **Assumption CL1:** $E[S_{i,j+1} \mid S_{i1}, S_{i2} \ldots, S_{ij}] = S_{ij} f_j$, independently of i for some set of parameters f_j (age-to-age factors)
- Note that this implies

$$E[Y_{ij}] = \alpha_i \beta_j$$

for parameters α_i, β_j

A slightly different model (cont'd)

Chain ladder model

$$E[S_{i,j+1} \mid S_{i1}, S_{i2}, \ldots, S_{ij}] = S_{ij} f_j$$

Y_{ij} stochastically independent as between **rows of triangle**

A slightly different model (cont'd)

Chain ladder model

$$E[S_{i,j+1} \mid S_{i1},S_{i2}\ldots,S_{ij}] = S_{ij}f_j$$

Y_{ij} stochastically independent as between **rows of triangle**

Cross-classified model

$$E[Y_{ij}] = \alpha_i\beta_j$$

Y_{ij} stochastically independent as between **all observations**

A slightly different model (cont'd)

Chain ladder model | **Cross-classified model**

$\mathbf{E[S_{i,j+1} \mid S_{i1},S_{i2}...,S_{ij}] = S_{ij}f_j}$ → $\mathbf{E[Y_{ij}] = \alpha_i\beta_j}$

Y_{ij} stochastically independent as between **rows of triangle** ← Y_{ij} stochastically independent as between **all observations**

Neither model more general than the other

Distribution of chain ladder observations

- We wish to investigate ML estimation for chain ladder model
- Need to specify likelihood of the Y_{ij}

Exponential dispersion family (EDF)

- Log-likelihood is

$$\ell(y;\theta,\lambda) = c(\lambda)[y\theta - b(\theta)] + a(y,\lambda)$$

  for some functions a(.,.), b(.) and c(.) and parameters θ and λ

- May be shown that

$$\mu = E[Y] = b'(\theta)$$

$$Var[Y] = b''(\theta)/c(\lambda)$$

Tweedie family of distributions

- EDF log-likelihood:

$$\ell(y;\theta,\lambda) = c(\lambda)[y\theta - b(\theta)] + a(y,\lambda)$$

- A subset of the EDF is obtained by means of the following restrictions:

$$c(\lambda) = \lambda$$

$$\text{Var}[Y] = \mu^p/\lambda,\ p \le 0 \text{ or } p \ge 1$$

- This restricts the log-likelihood to

$$\ell(y;\theta,\lambda) = \lambda[y\theta - b(\theta)] + a(y,\lambda)$$

with the 2nd restriction causing a restriction on the form of b(θ)

- The Tweedie subset of the EDF is the set of distributions used by most GLM regression packages (e.g. SAS PROC GENMOD)

Known members of the Tweedie family

$$\ell(y;\theta,\lambda) = \lambda[y\theta - b(\theta)] + a(y,\lambda)$$

$$\mathrm{Var}[Y] = \mu^p/\lambda$$

- Special cases

p	Distrib-ution	b(θ)
0	Normal	$\frac{1}{2}\theta^2$
1	Poisson	$\exp\theta$
2	Gamma	$-\ln(-\theta)$
3	Inverse Gaussian	$-(-2\theta)^{1/2}$
(1,2)	Compound Poisson - gamma	

Known members of the Tweedie family

$$\ell(y;\theta,\lambda) = \lambda[y\theta - b(\theta)] + a(y,\lambda)$$

$$Var[Y] = \mu^p/\lambda$$

- Special cases

NOTE: For case p=1

$$Var[Y] = \mu/\lambda$$

which is more general than Poisson ($Var[Y] = \mu$)

This distribution is called **over-dispersed Poisson (ODP)**

p	**Distrib -ution**	**$b(\theta)$**
0	Normal	$\frac{1}{2}\theta^2$
1	Poisson	$\exp\theta$
2	Gamma	$-\ln(-\theta)$
3	Inverse Gaussian	$-(-2\theta)^{\frac{1}{2}}$
(1,2)	Compound Poisson - gamma	

MLE for Tweedie data

- Model form
  - Cross-classified model with Tweedie distributed observations
  - Slightly generalised variance

$$\mathrm{Var}\,[Y_{ij}] = \mu_{ij}^{p}/\lambda w_{ij}$$

where w_{ij} is new and is the **weight** associated with Y_{ij}

MLE equations are:

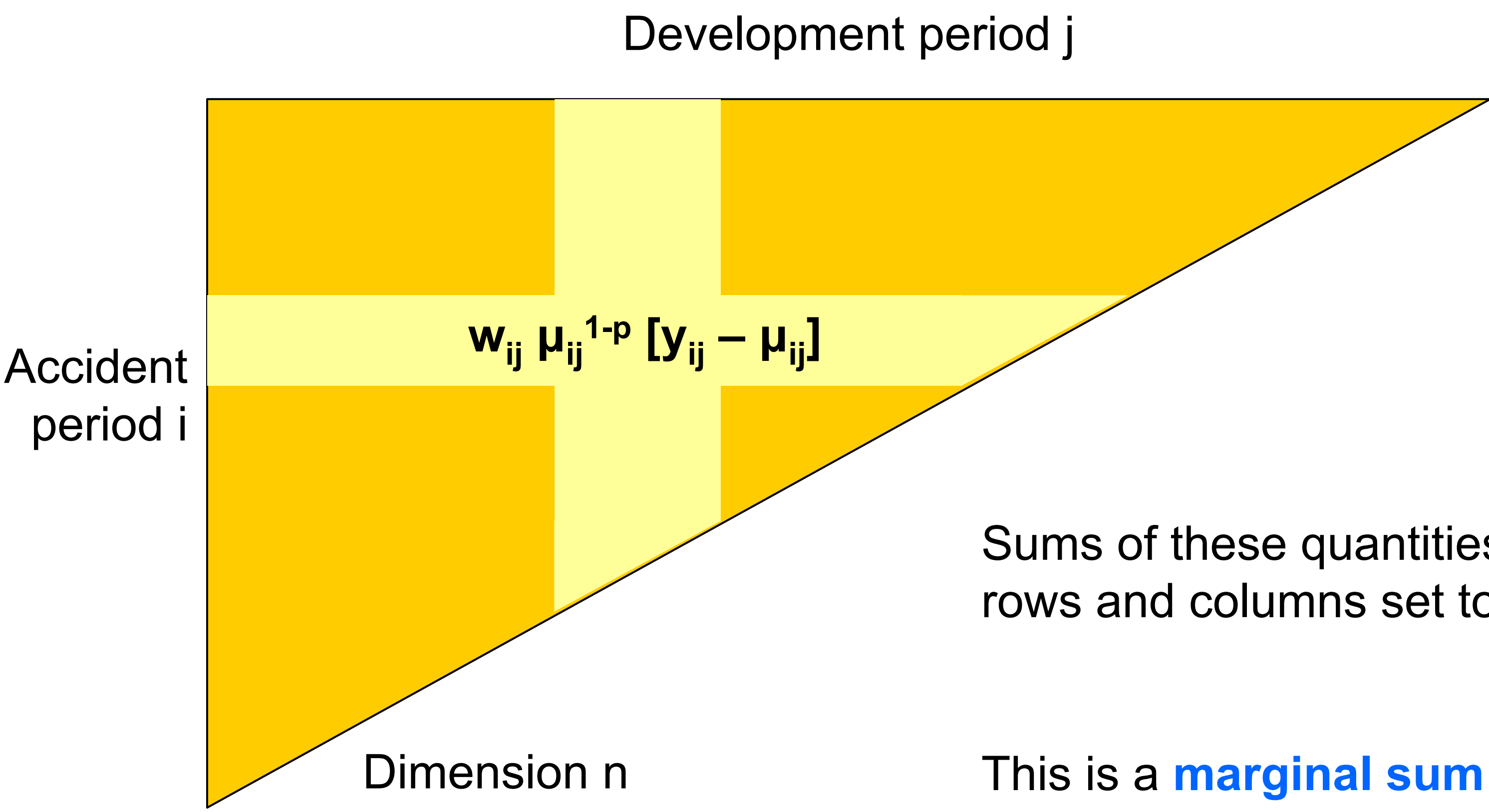
$$\sum^{R(i)} w_{ij}\, \mu_{ij}^{1-p} [y_{ij} - \mu_{ij}] = 0,\ i=1,\ldots,n$$
$$\sum^{C(j)} w_{ij}\, \mu_{ij}^{1-p} [y_{ij} - \mu_{ij}] = 0,\ j=1,\ldots,n$$

where

$\sum^{R(i)}$ denotes summation over the entire row i of the triangle

$\sum^{C(j)}$ denotes summation over the entire column j

MLE for Tweedie data (cont'd)



Sums of these quantities over all rows and columns set to zero

This is a **marginal sum** method of estimation

MLE for Tweedie data – Special case 1

- MLE equations

$$\sum^{R(i)} w_{ij}\, \mu_{ij}^{1-p}\, [y_{ij} - \mu_{ij}] = 0,\ i=1,\ldots,n$$

$$\sum^{C(j)} w_{ij}\, \mu_{ij}^{1-p}\, [y_{ij} - \mu_{ij}] = 0,\ j=1,\ldots,n$$

- Special case **$p=w_{ij}=1$** (over-dispersed Poisson Y_{ij})

$$\sum^{R(i)} [y_{ij} - \mu_{ij}] = 0,\ i=1,\ldots,n$$

$$\sum^{C(j)} [y_{ij} - \mu_{ij}] = 0,\ j=1,\ldots,n$$

- The solution to this system is known to be **chain ladder** (Hachemeister & Stanard, 1975)

MLE for Tweedie data – Special case 2

- MLE equations

$$\sum^{R(i)} w_{ij}\, \mu_{ij}^{1-p}\, [y_{ij} - \mu_{ij}] = 0,\ i=1,\ldots,n$$
$$\sum^{C(j)} w_{ij}\, \mu_{ij}^{1-p}\, [y_{ij} - \mu_{ij}] = 0,\ j=1,\ldots,n$$

- Special case **p=2** (gamma Y_{ij})

$$\sum^{R(i)} w_{ij}\, [y_{ij} / \mu_{ij} - 1] = 0,\ i=1,\ldots,n$$
$$\sum^{C(j)} w_{ij}\, [y_{ij} / \mu_{ij} - 1] = 0,\ j=1,\ldots,n$$

- The solution to this system was studied by Mack (1991)

MLE for Tweedie data – chain ladder as a limiting case

- It may be shown that the chain ladder approximates the solution to the Tweedie cross-classified model if any of the following conditions holds:
  - Observation variances are small
  - p is close to 1
    - compound (over-dispersed) Poisson with gamma severity with low coefficient of variation)
  - Weights $w_{ij}\ \mu_{ij}^{1-p}$ vary little over the triangle

MLE for Tweedie data – multiplicative weights

- MLE equations

$$\sum^{R(i)} w_{ij}\, \mu_{ij}^{1-p}\, [y_{ij} - \mu_{ij}] = 0,\ i=1,\ldots,n$$
$$\sum^{C(j)} w_{ij}\, \mu_{ij}^{1-p}\, [y_{ij} - \mu_{ij}] = 0,\ j=1,\ldots,n$$

- Consider case of multiplicative weights

$$w_{ij} = u_i\, v_j$$

- Recall

$$\mu_{ij} = \alpha_i \beta_j$$

MLE for Tweedie data – multiplicative weights

- MLE equations

$$\sum^{R(i)} w_{ij}\, \mu_{ij}^{1-p}\, [y_{ij} - \mu_{ij}] = 0,\ i=1,\ldots,n$$
$$\sum^{C(j)} w_{ij}\, \mu_{ij}^{1-p}\, [y_{ij} - \mu_{ij}] = 0,\ j=1,\ldots,n$$

- Consider case of multiplicative weights

$$w_{ij} = u_i\, v_j$$

- Recall

$$\mu_{ij} = \alpha_i \beta_j$$

- Hence

$$\sum^{R(i)} [w_{ij}\, \mu_{ij}^{1-p}\, y_{ij} - w_{ij}\, \mu_{ij}^{2-p}] = 0,\ i=1,\ldots,n$$
$$\sum^{C(j)} [w_{ij}\, \mu_{ij}^{1-p}\, y_{ij} - w_{ij}\, \mu_{ij}^{2-p}] = 0,\ j=1,\ldots,n$$

$$\sum^{R(i)} [z_{ij} - \xi_{ij}] = 0,\ i=1,\ldots,n$$
$$\sum^{C(j)} [z_{ij} - \xi_{ij}] = 0,\ j=1,\ldots,n$$

with

$$Z_{ij} = w_{ij}\, \mu_{ij}^{1-p}\, Y_{ij}$$
$$\xi_{ij} = w_{ij}\, \mu_{ij}^{2-p} = [u_i \alpha_i^{2-p}]\, [v_j \beta_j^{2-p}] = a_i\, b_j$$

MLE for Tweedie data – multiplicative weights

$$\sum^{R(i)} [z_{ij} – \xi_{ij}] = 0, i=1,\ldots,n$$

$$\sum^{C(j)} [z_{ij} – \xi_{ij}] = 0, j=1,\ldots,n$$

with

$$Z_{ij} = w_{ij} \mu_{ij}^{1-p} Y_{ij}$$

$$\xi_{ij} = [u_i\alpha_i^{2-p}] [v_j\beta_j^{2-p}] = a_i b_j$$

- Note that these are chain ladder equations for observations Z_{ij} and parameters a_i, b_j
- But solution of chain ladder requires fore-knowledge of μ_{ij} which are targets of estimation
- Solution proceeds by iteration

MLE for Tweedie data – numerical solution for multiplicative weights

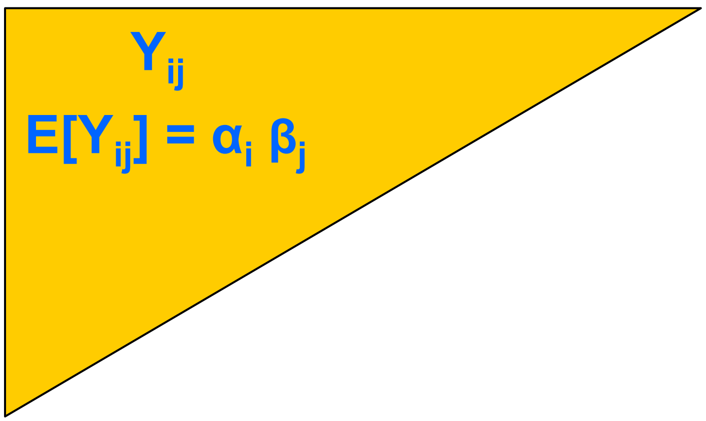
- Chain ladder on
  - Observations $Z_{ij} = w_{ij}\, \mu_{ij}^{1-p}\, Y_{ij}$ $[\mu_{ij} = \alpha_i \beta_j]$
  - With expectations $\xi_{ij} = [u_i \alpha_i^{2-p}]\, [v_j \beta_j^{2-p}] = a_i\, b_j$
- In r-th iteration, apply chain ladder estimation to
  - Observations $Z^{(r)}_{ij} = w_{ij}\, (\mu^{(r)}_{ij})^{1-p}\, Y_{ij}$ $[\mu^{(r)}_{ij} = \alpha^{(r)}_i \beta^{(r)}_j]$
  - With expectations $\xi^{(r)}_{ij} =$

    $$[u_i (\alpha^{(r+1)}_i)^{2-p}]\, [v_j (\beta^{(r+1)}_j)^{2-p}] = a^{(r+1)}_i\, b^{(r+1)}_j$$
- This produces new estimates $a^{(r+1)}_i$, $b^{(r+1)}_j$

MLE for Tweedie data – numerical solution for multiplicative weights

Diagrammatically

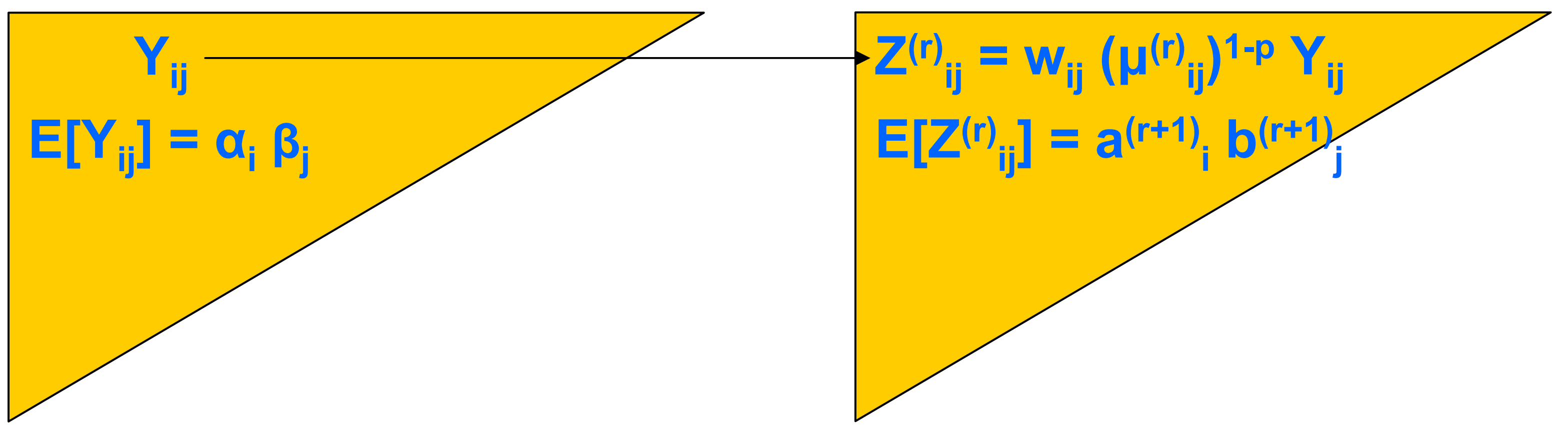
Actual variables



MLE for Tweedie data – numerical solution for multiplicative weights

Diagrammatically

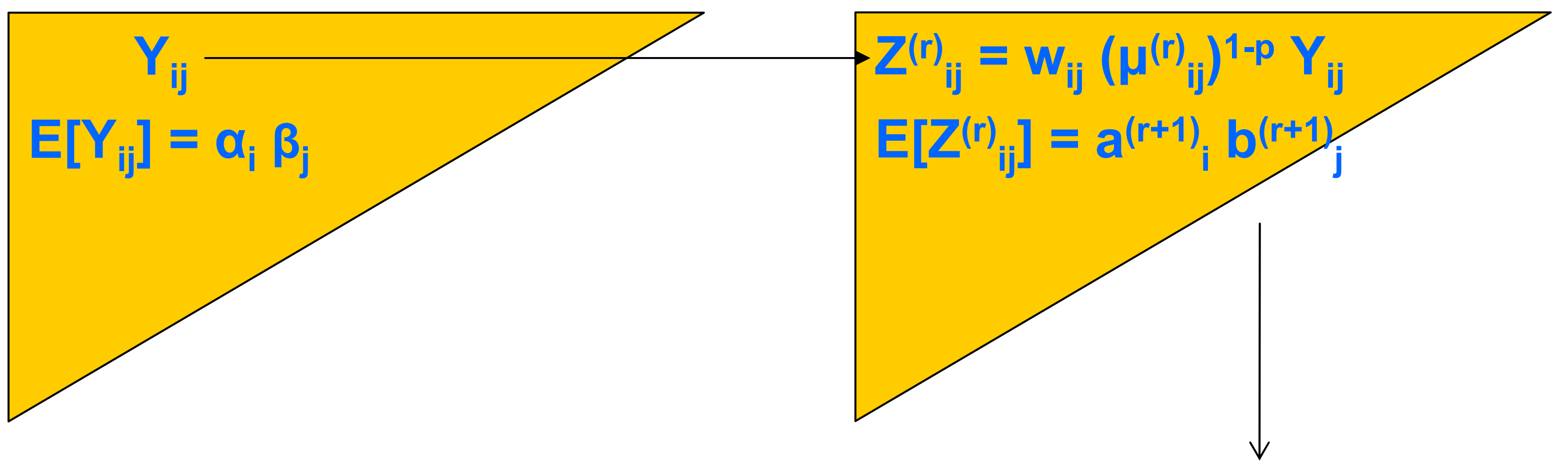
Actual variables **Transformed variables**



MLE for Tweedie data – numerical solution for multiplicative weights

Diagrammatically

Actual variables **Transformed variables**

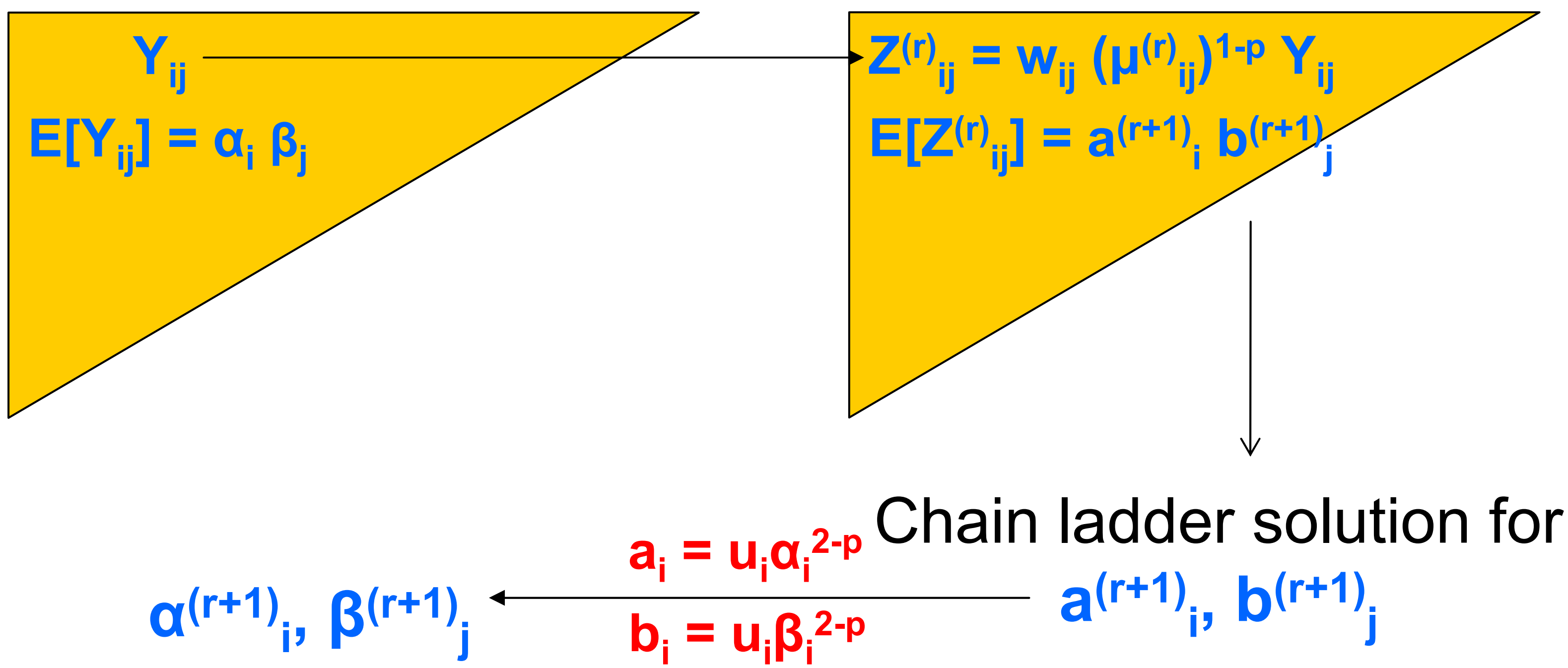


Chain ladder solution for
$a^{(r+1)}_i, b^{(r+1)}_j$

MLE for Tweedie data – numerical solution for multiplicative weights

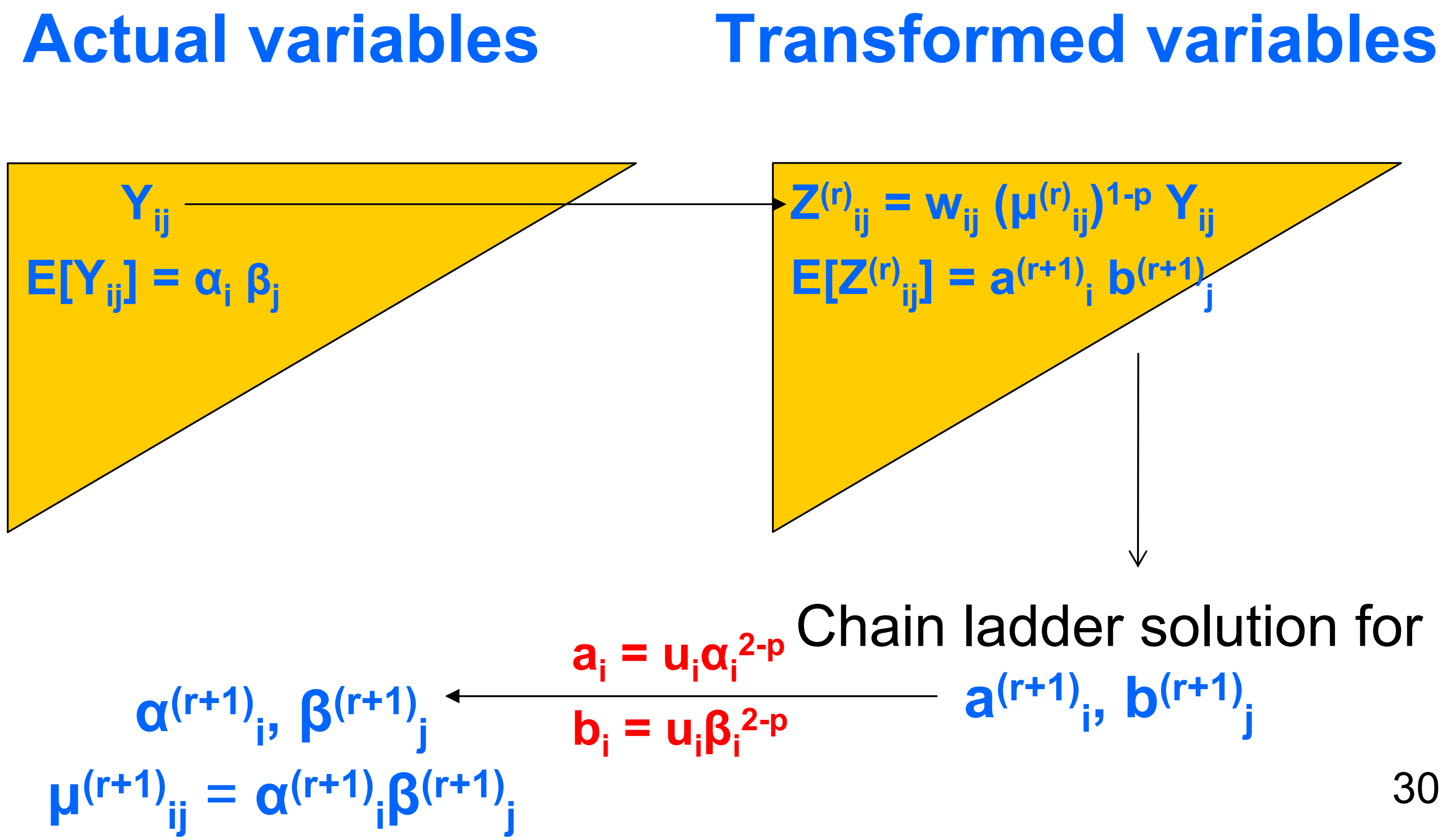
Diagrammatically

Actual variables **Transformed variables**



MLE for Tweedie data – numerical solution for multiplicative weights

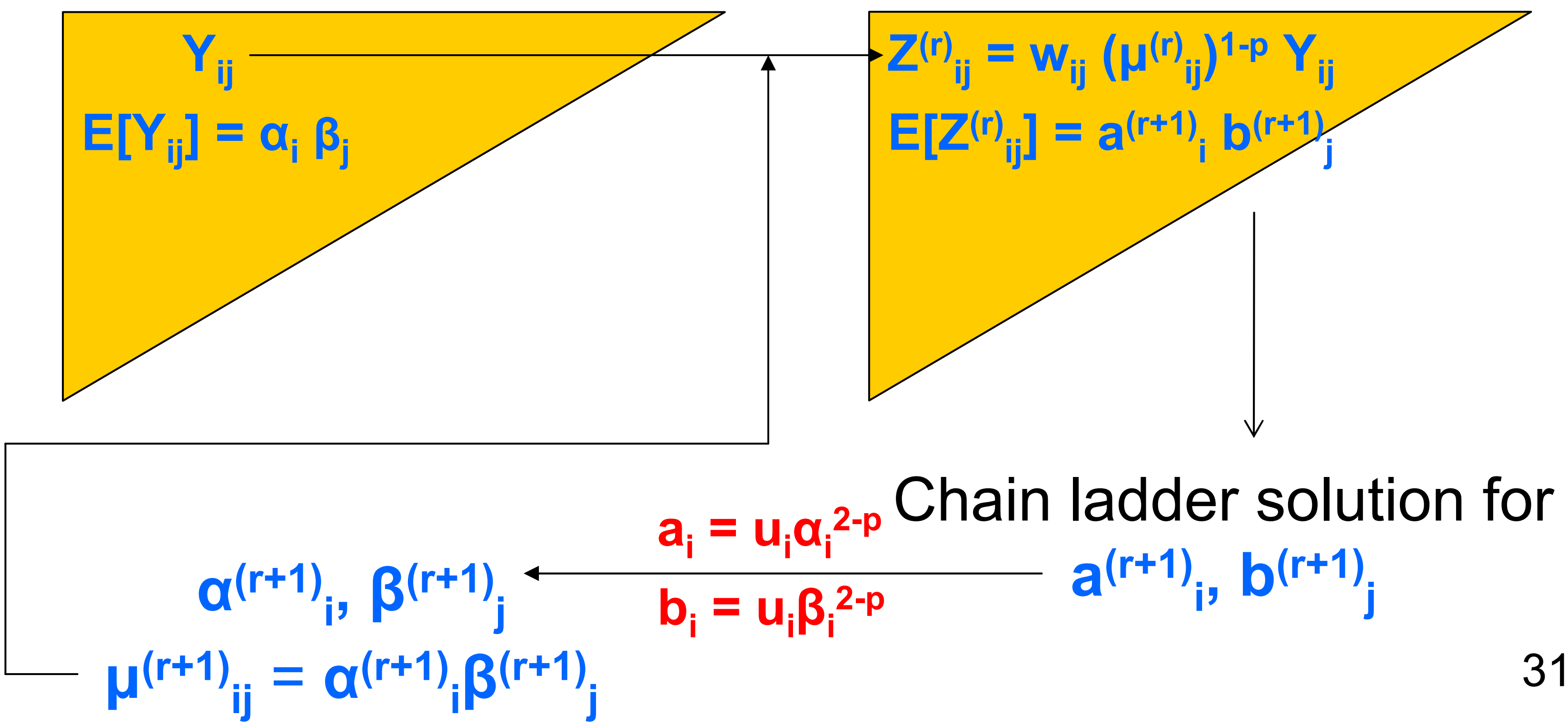
Diagrammatically



MLE for Tweedie data – numerical solution for multiplicative weights

Diagrammatically

Actual variables **Transformed variables**



MLE for Tweedie data – numerical solution for multiplicative weights

- This iterative procedure converges quickly for small values of p
  - Shorter tailed distributions
  - Converges in a single iteration for p=1
- Converges more slowly as p increases
- Recall that 1<p<2 for compound Poisson observations
- Numerical experiment requiring convergence to accuracy of 0.05% in loss reserve:
  - p=2 (gamma): 5 iterations
  - p=2.4 (fairly long tailed): 24 iterations

Practical implications

- The chain ladder (with multiplicative weights $w_{ij}=u_i v_j$) will give MLE for p=1
- For compound Poisson cells of the triangle $1<p<2$, and the chain ladder is **not** MLE
- The difference from MLE increases with p
- Larger p means more extreme variance for large mean values ($Var[Y_{ij}] = \mu^p/\lambda$)
- In practice one needs to consider the likely value of p in this relation in deciding whether or not the chain ladder is likely to produce efficient estimates (and forecasts)