General
Insurance
Pricing
Seminar

Institute of Actuaries of Australia

Computation of credibility coefficients for pricing

Greg Taylor

Overview

- Statement of the problem
- Fundamentals of credibility theory
- Estimation of credibility coefficients in simple models
- Analysis of variance
- Extension to more general models

Overview

- Material taken from:

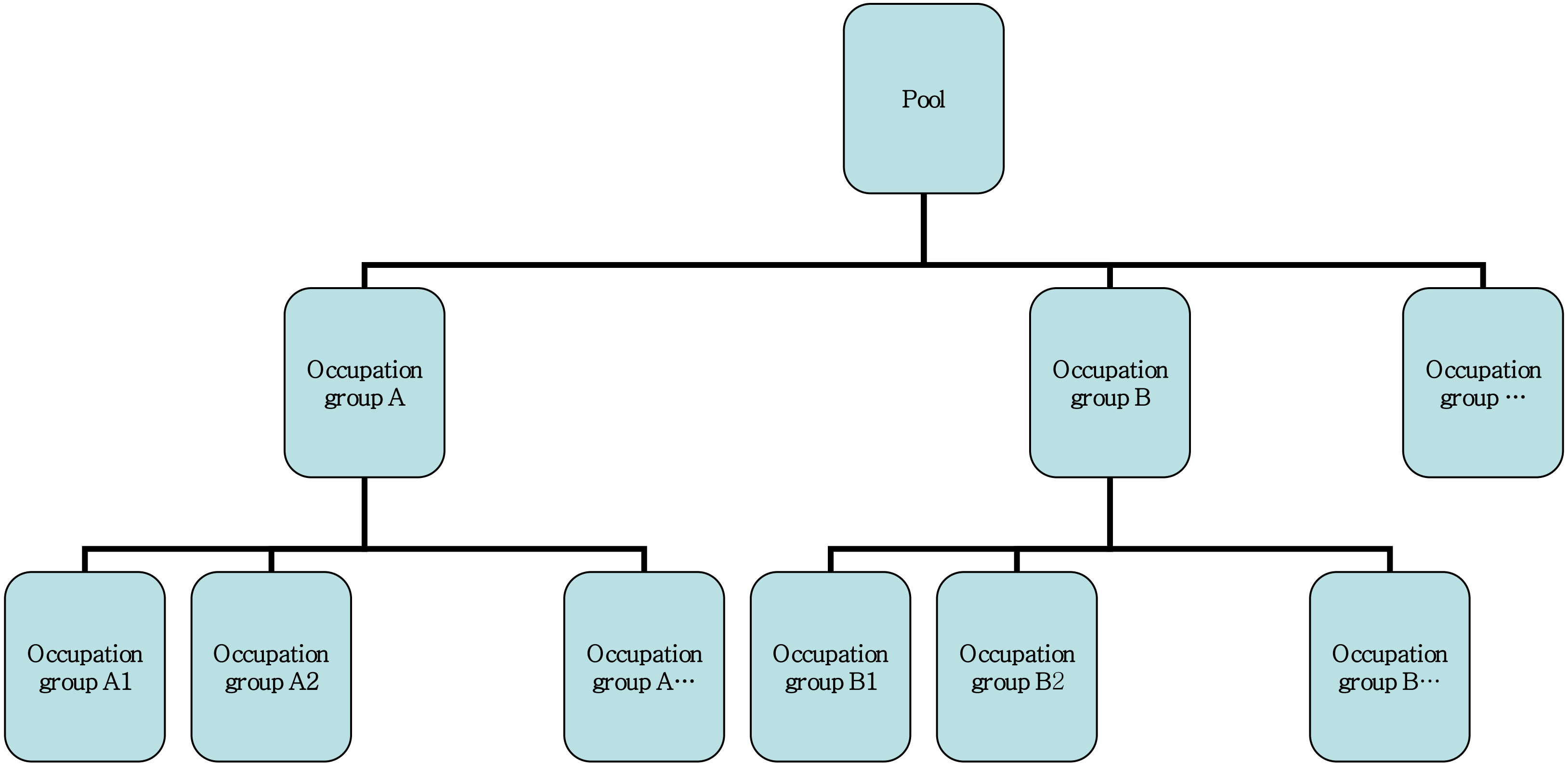
Taylor G (2007). Credibility, hypothesis testing and regression software. **Astin Bulletin**, 37 (in press)

- Also appears as University of Melbourne Research Paper No. 149 at http://www.economics.unimelb.edu.au/SITE/actwww/wps2007/No149.pdf

Statement of the problem

- Suppose you wish to estimate some pricing parameter (e.g. claim frequency)
- You have a measurement of it from data
  - but subject to sampling error
- You also have some prior information on it from somewhere (e.g. industry data)
  - but also subject to uncertainty
- You wish to form an estimate of the parameter that takes both pieces of information into account
- How should you weight those two pieces of information?

Example 1 – workers compensation rating by ANZSIC occupation code (Taylor, 1979)



etc.

Example 2 – multiplicative pricing structure (e.g. Motor) (Gisler & Müller, 2007)

- Usual multiplicative pricing structure

		Expected values			
		Pricing factor A =			
		1	**2**	**…**	**J**
Pricing factor B =	**1**	$\alpha_1\beta_1$	$\alpha_1\beta_2$	…	$\alpha_1\beta_J$
	2	$\alpha_2\beta_1$	$\alpha_2\beta_2$	…	$\alpha_2\beta_J$
	:				
	K	$\alpha_K\beta_1$	$\alpha_K\beta_2$	…	$\alpha_K\beta_J$

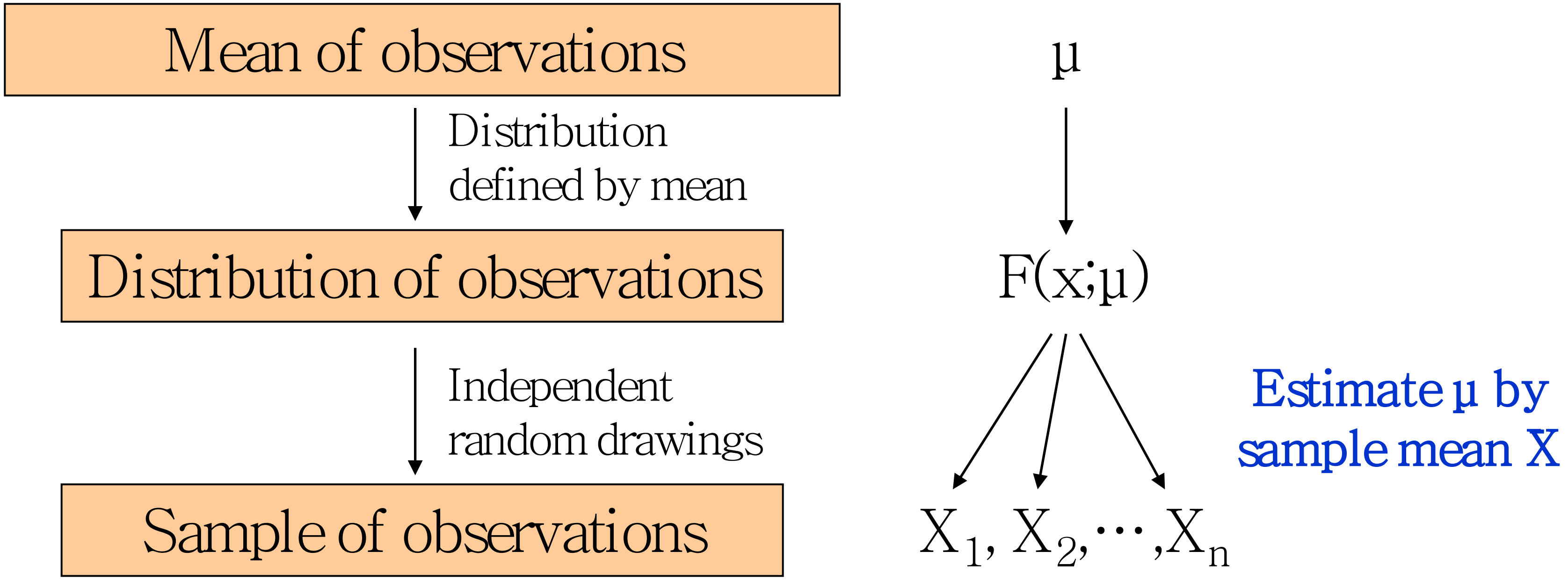
Example 2 – multiplicative pricing structure (e.g. Motor) (Gisler & Müller, 2007)

- Usual multiplicative pricing structure
- But suppose that data is sparse for some values of a pricing factor
  - E.g. no recorded claims for pricing factor A=2
  - GLM will generate fitted values of zero for A=2
- How might model be changed to give reasonable results in this case?

		Expected values			
		Pricing factor A =			
		1	**2**	**…**	**J**
Pricing factor B =	**1**	$\alpha_1\beta_1$	$\alpha_1\beta_2$	…	$\alpha_1\beta_J$
	2	$\alpha_2\beta_1$	$\alpha_2\beta_2$	…	$\alpha_2\beta_J$
	:				
	K	$\alpha_K\beta_1$	$\alpha_K\beta_2$	…	$\alpha_K\beta_J$

Fundamentals of credibility theory

Sampling a population



Mean of observations

Distribution defined by mean

Distribution of observations

Independent random drawings

Sample of observations

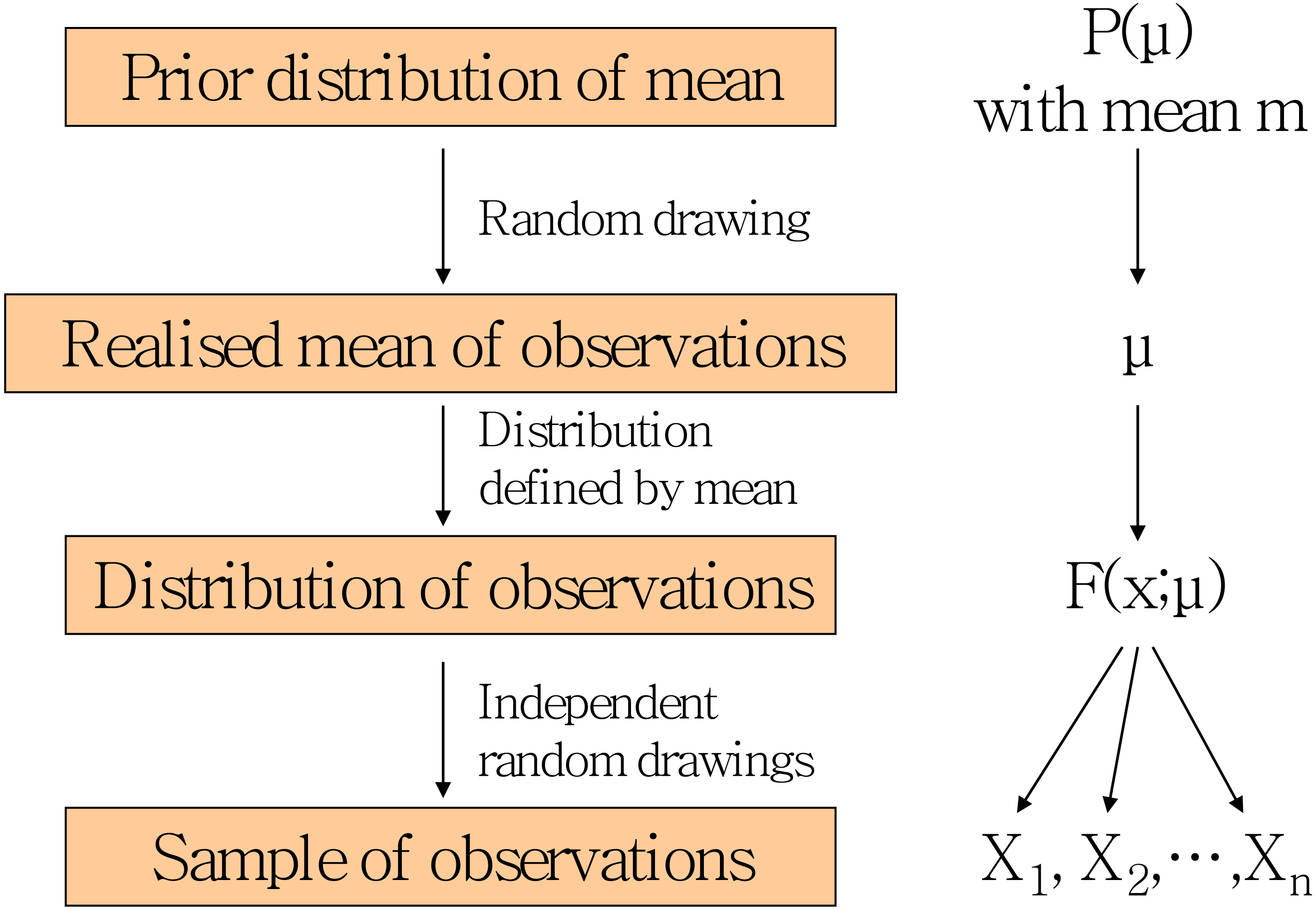
μ

$F(x;\mu)$

$X_1, X_2, \cdots, X_n$

Estimate μ by sample mean X

Sampling a population with a random mean



Prior distribution of mean

Random drawing

Realised mean of observations

Distribution defined by mean

Distribution of observations

Independent random drawings

Sample of observations

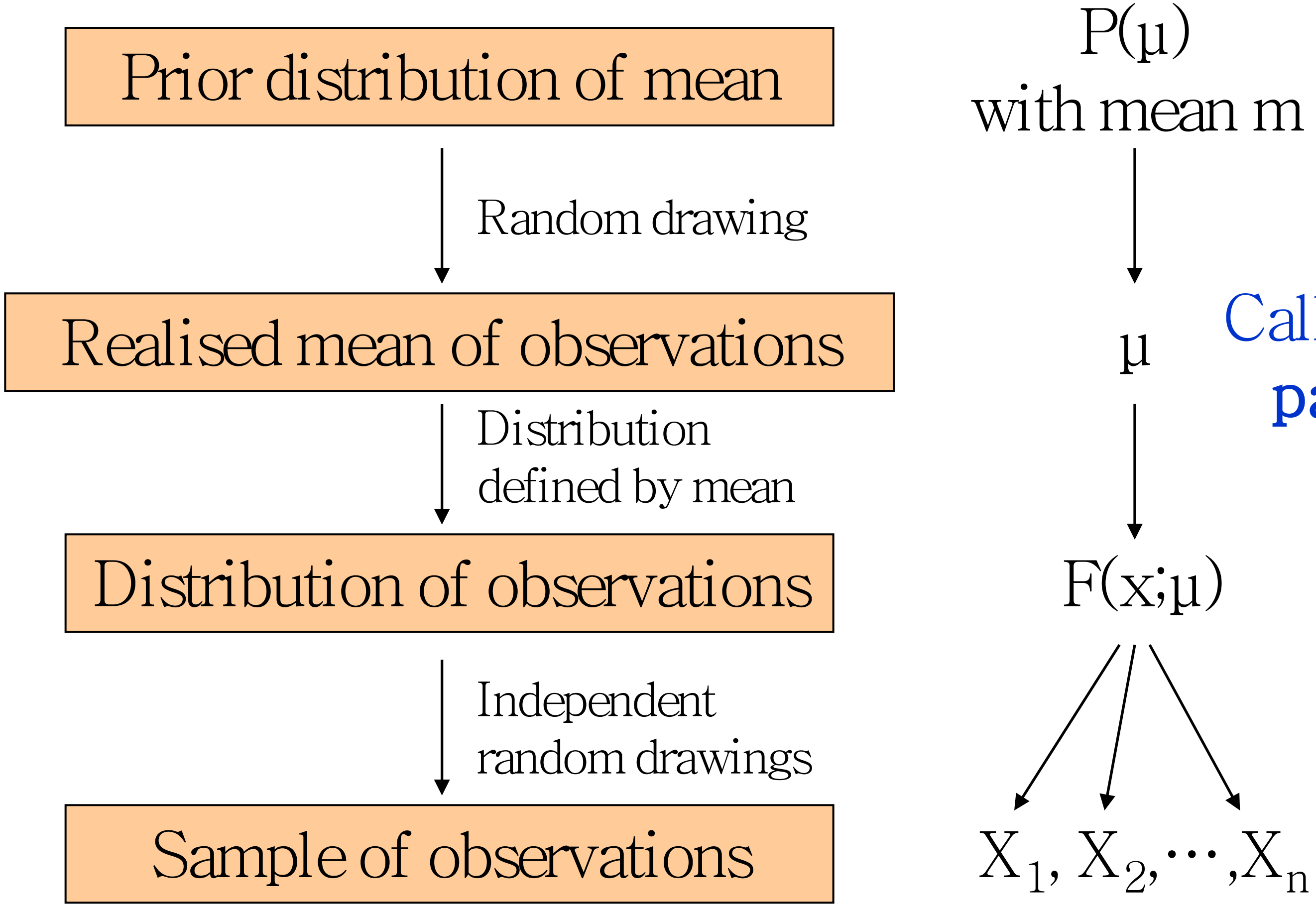
$P(\mu)$ with mean m

μ

$F(x;\mu)$

$X_1, X_2, \cdots, X_n$

Sampling a population with a random mean



Prior distribution of mean

Random drawing

Realised mean of observations

Distribution defined by mean

Distribution of observations

Independent random drawings

Sample of observations

$P(\mu)$ with mean m

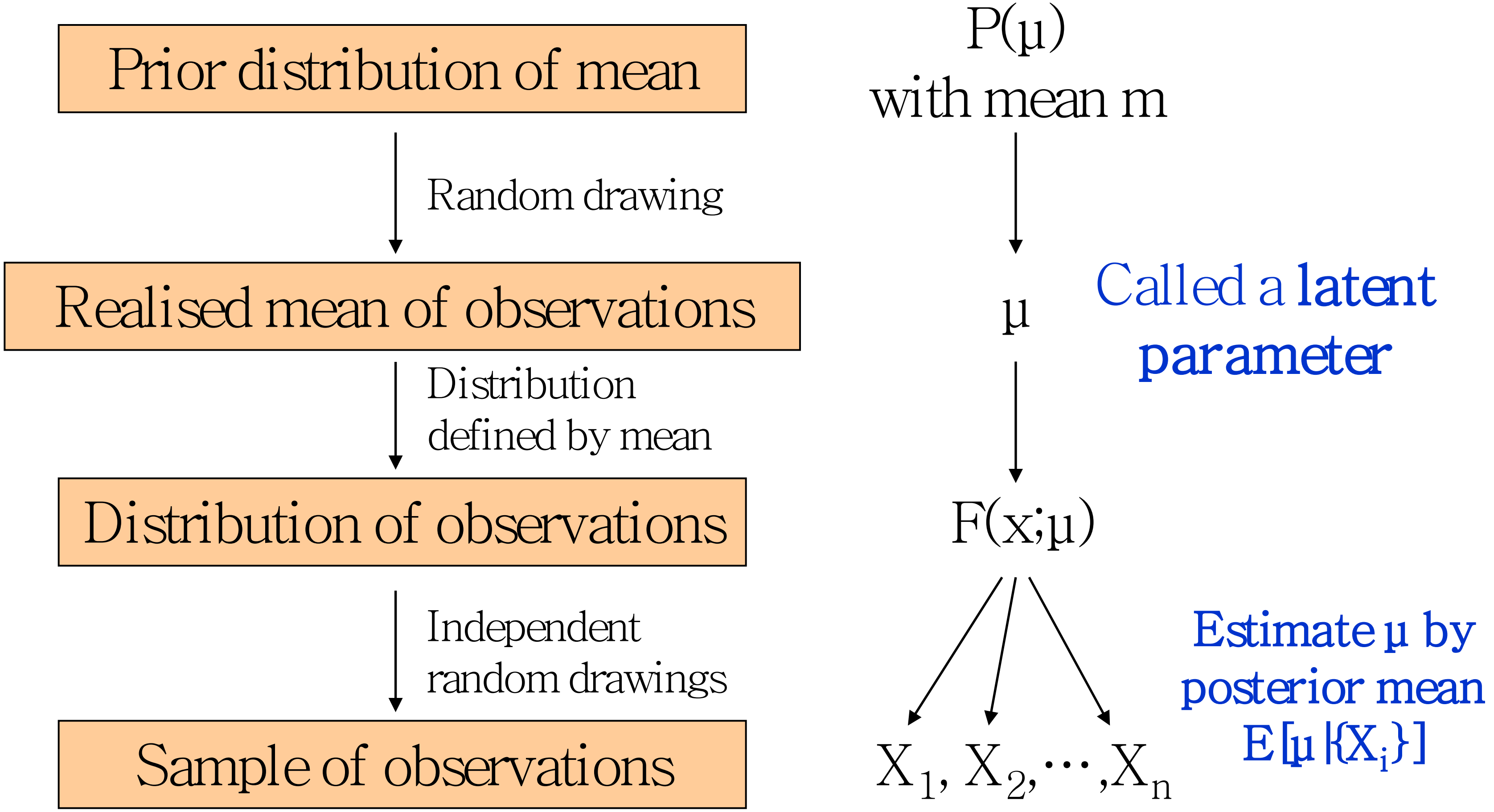
μ

Called a **latent parameter**

$F(x;\mu)$

$X_1, X_2, \cdots, X_n$

Estimation of a random mean



Prior distribution of mean

Random drawing

Realised mean of observations

Distribution defined by mean

Distribution of observations

Independent random drawings

Sample of observations

$P(\mu)$ with mean m

μ

Called a **latent parameter**

$F(x;\mu)$

$X_1, X_2, \cdots, X_n$

Estimate μ by posterior mean $E[\mu|\{X_i\}]$

Bayesian framework

Let $X=(X_1, X_2,\cdots,X_n)$

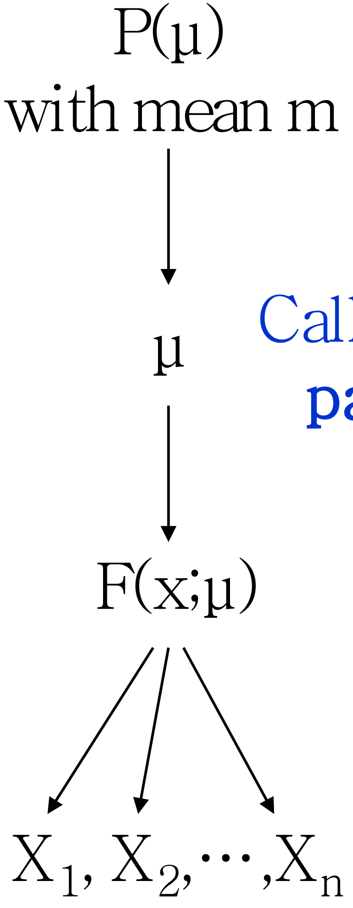
By Bayes theorem

$$E[\mu|X] = \int \mu\, p(\mu|X)\, d\mu$$

$$= \frac{\int \mu\, dP(\mu)\, p(X|\mu)}{\int dP(\mu)\, p(X|\mu)}$$

$$= \frac{\int \mu\, dP(\mu) \int dF(X_1|\mu)\cdots dF(X_n|\mu)}{\int dP(\mu) \int dF(X_1|\mu)\cdots dF(X_n|\mu)}$$

Estimate $E[\mu|X]$ by a linear function $L(X)$ of X


$P(\mu)$
with mean m

μ

$F(x;\mu)$

$X_1, X_2,\cdots,X_n$

Called a **latent parameter**

Estimate μ by posterior mean $E[\mu|\{X_i\}]$

Linear Bayes framework

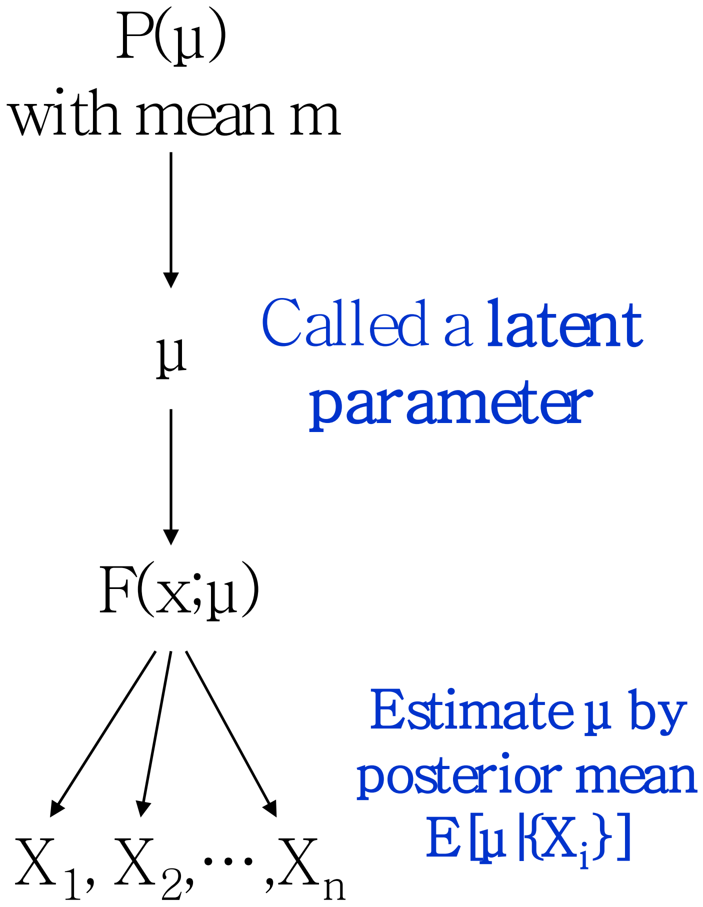
Estimate $E[\mu|X]$ by a linear function $L(X)$ of X

$L(X)$ is called a **linear Bayes estimator**

Choose so as to minimise

$$\int [L(X) - \mu]^2 \, p(\mu,X) \, d\mu \, dX$$

$$= \int [L(X) - \mu]^2 \, dP(\mu) \int dF(X_1|\mu) \cdots dF(X_n|\mu)$$



$P(\mu)$

with mean m

μ

Called a **latent parameter**

$F(x;\mu)$

$X_1, X_2, \cdots, X_n$

Estimate μ by posterior mean $E[\mu|\{X_i\}]$

Linear Bayes framework

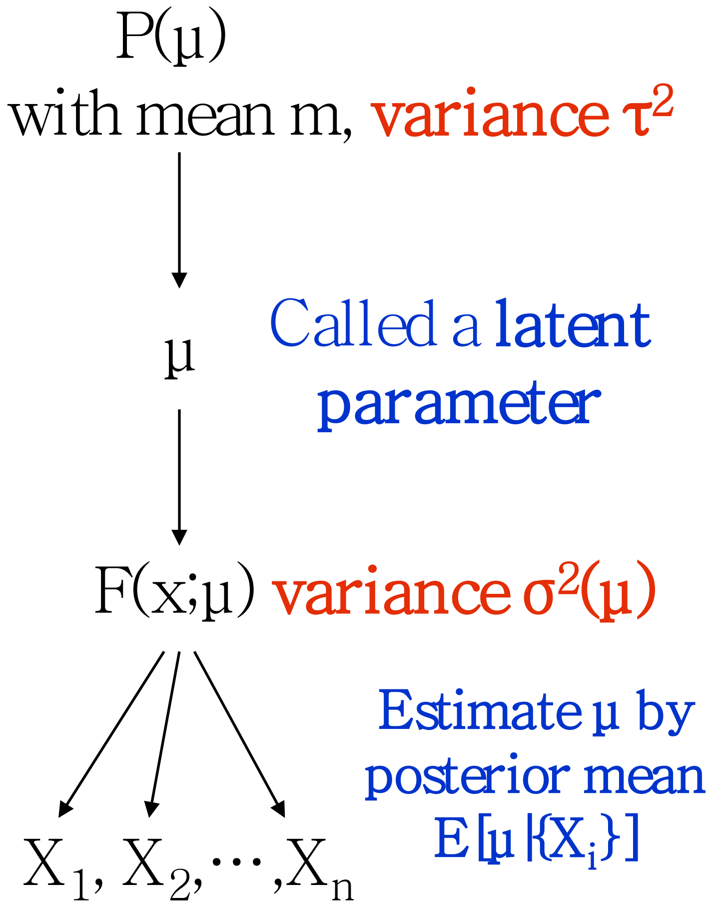
Estimate $E[\mu | X]$ by a linear function $L(X)$ of X

It may be shown that

$$L(X) = (1-z)\underset{\substack{\uparrow \\ \text{Prior mean}}}{m} + \underset{\substack{\uparrow \\ \text{Credibility of } \bar{X}}}{z}\,\underset{\substack{\uparrow \\ \text{Data mean}}}{\bar{X}}$$

where

$$z = \{1 + E_\mu[\underset{\substack{\uparrow \\ \text{Data variance}}}{\sigma^2(\mu)}]/n\underset{\substack{\uparrow \\ \text{Prior variance}}}{\tau^2}\}^{-1}$$



$P(\mu)$

with mean m, variance τ^2

μ

Called a **latent parameter**

$F(x;\mu)$ variance $\sigma^2(\mu)$

$X_1, X_2, \cdots, X_n$

Estimate μ by posterior mean $E[\mu | \{X_i\}]$

Linear Bayes framework

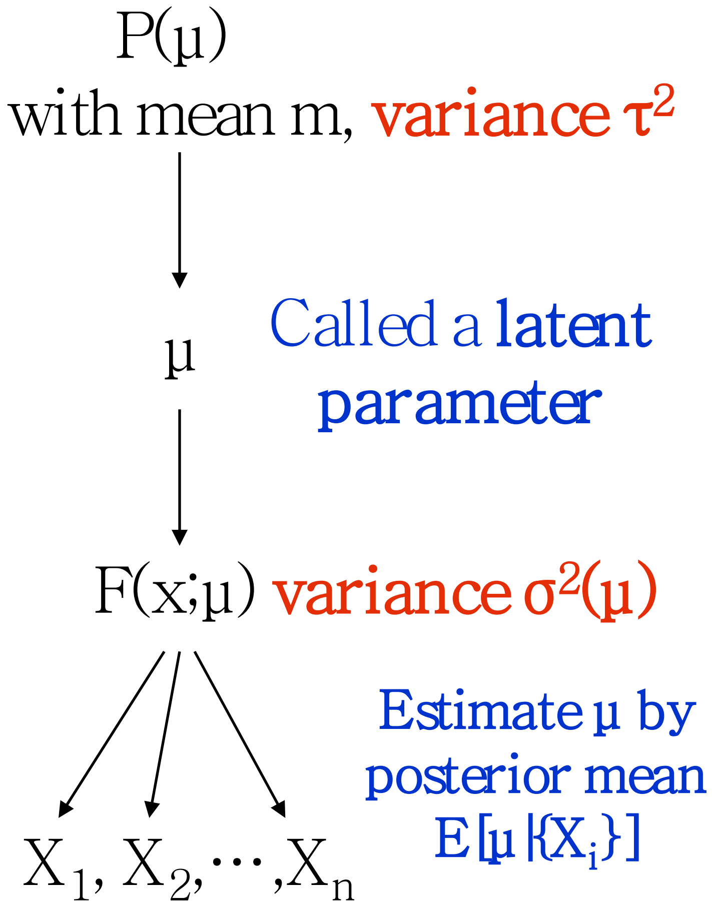
Credibility coefficient

$z = \{1 + E_\mu[\sigma^2(\mu)]/n\tau^2\}^{-1}$

There is a need to estimate the ratio of data variance to prior variance:

$E_\mu[\sigma^2(\mu)]/\tau^2$

To do so requires more data



$P(\mu)$

with mean m, variance τ^2

μ

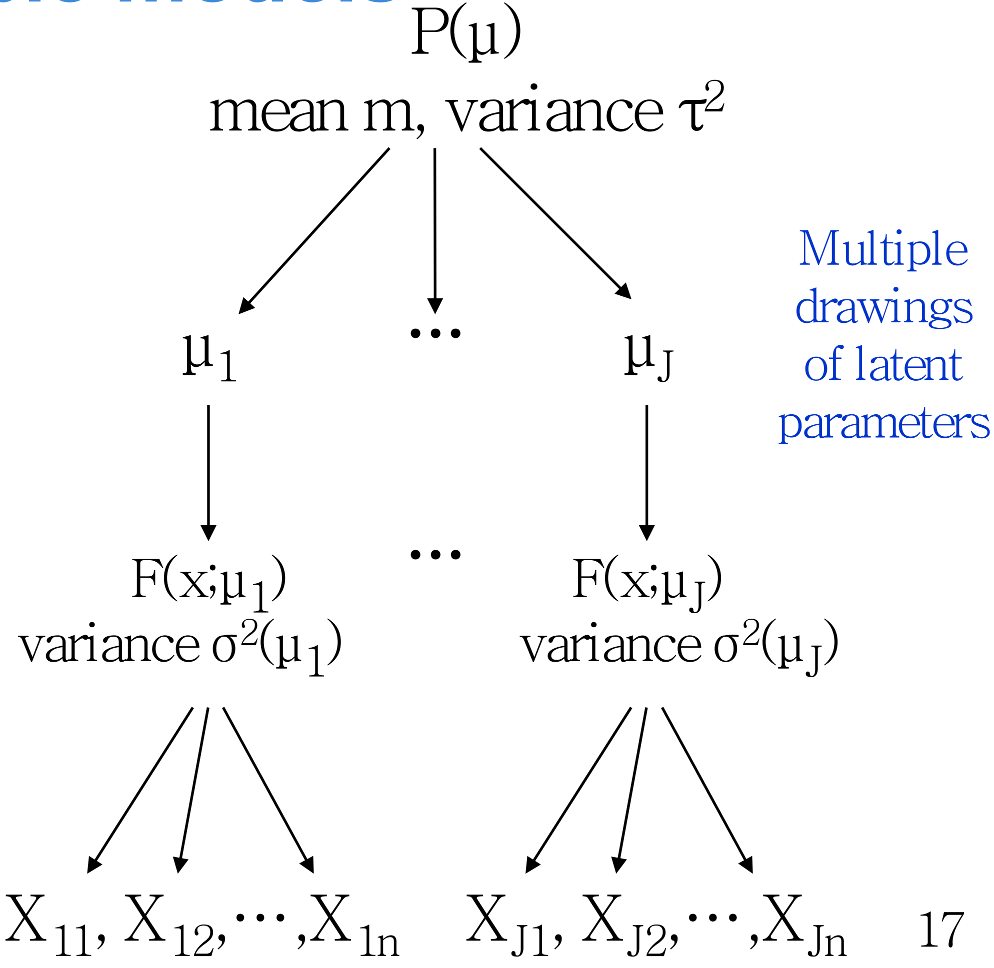
Called a **latent** parameter

$F(x;\mu)$ variance $\sigma^2(\mu)$

$X_1, X_2, \cdots, X_n$

Estimate μ by posterior mean $E[\mu|\{X_i\}]$

Estimation of credibility coefficients in simple models



$P(\mu)$

mean m, variance τ^2

μ_1 … μ_J

Multiple drawings of latent parameters

$F(x;\mu_1)$ … $F(x;\mu_J)$

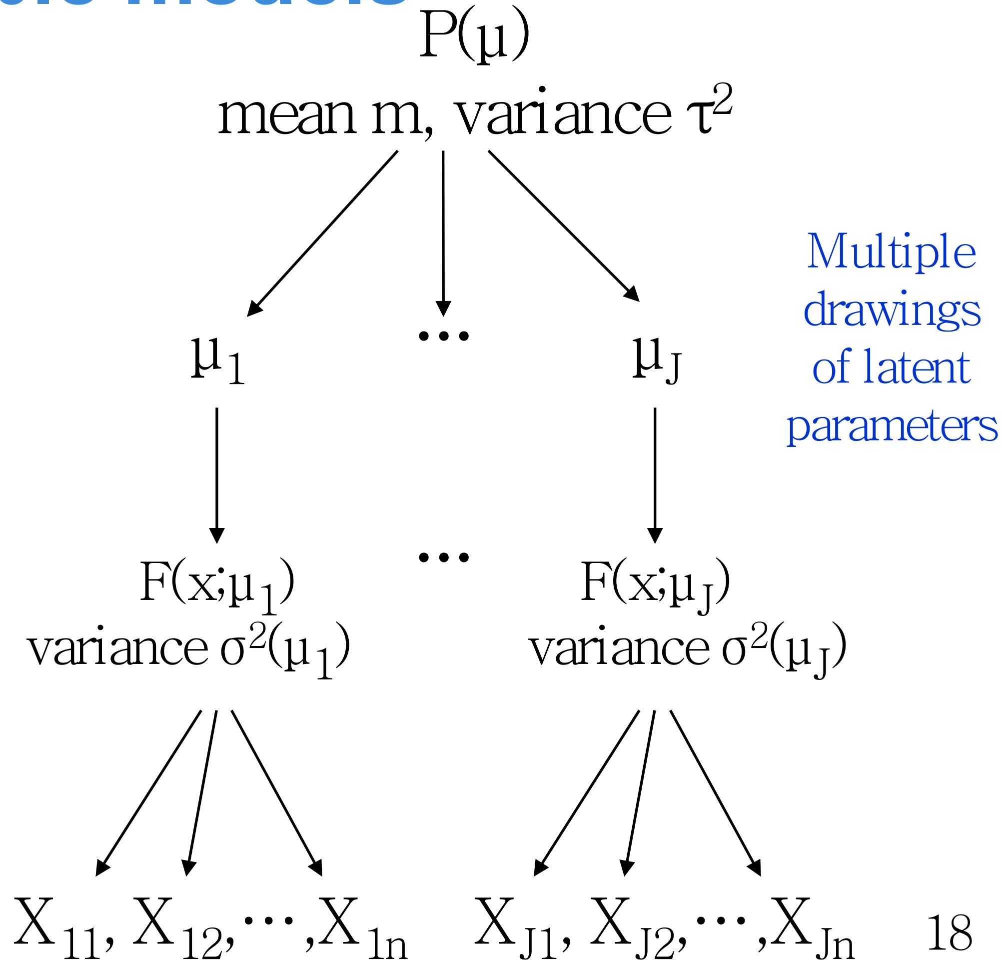
variance $\sigma^2(\mu_1)$ variance $\sigma^2(\mu_J)$

$X_{11}, X_{12}, \cdots, X_{1n}$ $X_{J1}, X_{J2}, \cdots, X_{Jn}$

Estimation of credibility coefficients in simple models

Sample $\mu_1, \mu_2, \cdots \mu_J$ independently from prior (J risk classes)

For each risk class μ_j, draw iid sample of n observations $X_{j1}, X_{j2}, \cdots, X_{jn}$



$P(\mu)$

mean m, variance τ^2

μ_1 … μ_J

$F(x;\mu_1)$ variance $\sigma^2(\mu_1)$ … $F(x;\mu_J)$ variance $\sigma^2(\mu_J)$

$X_{11}, X_{12}, \cdots, X_{1n}$ $X_{J1}, X_{J2}, \cdots, X_{Jn}$

Multiple drawings of latent parameters

Estimation of credibility coefficients in simple models

Sample $\mu_1, \mu_2, \cdots \mu_J$ from prior (J risk classes)

For each risk class μ_j, draw a iid sample of n observations $X_{j1}, X_{j2}, \cdots, X_{jn}$

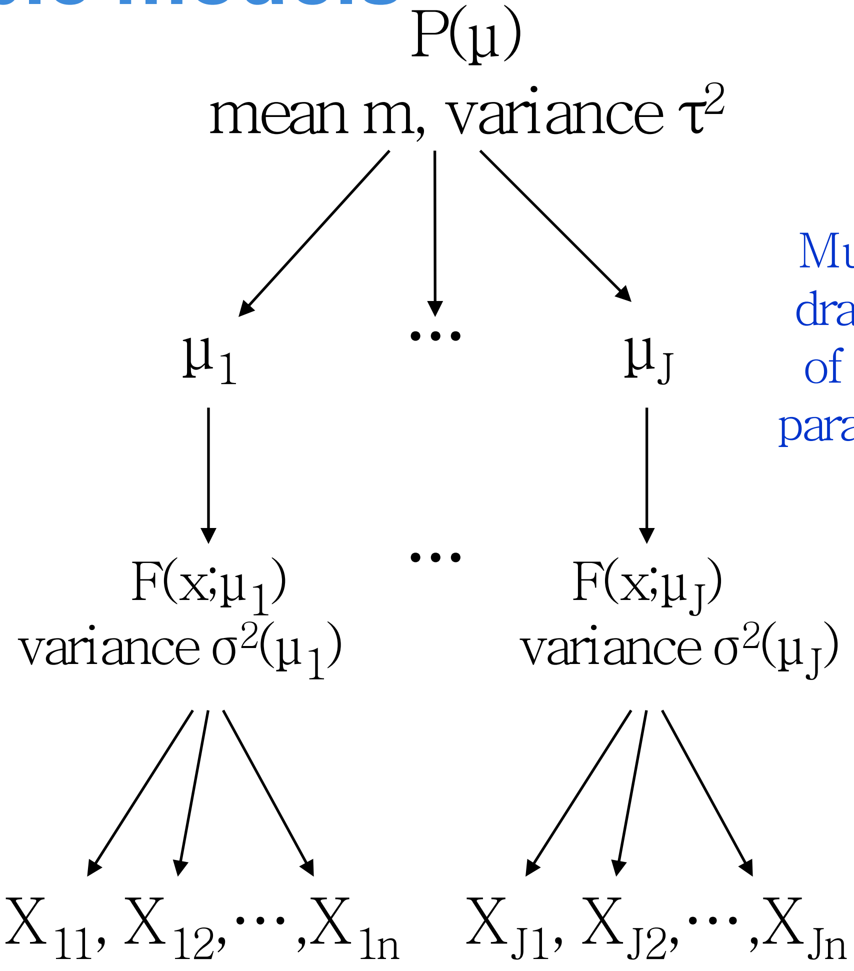
As before, estimate μ_j by

$$L(X) = (1-z)m + z\,\overline{X}_j$$

with z as before

Still need to estimate

$$E_\mu[\sigma^2(\mu)]/\tau^2$$



P(μ)

mean m, variance τ^2

μ_1 … μ_J

$F(x;\mu_1)$ variance $\sigma^2(\mu_1)$ … $F(x;\mu_J)$ variance $\sigma^2(\mu_J)$

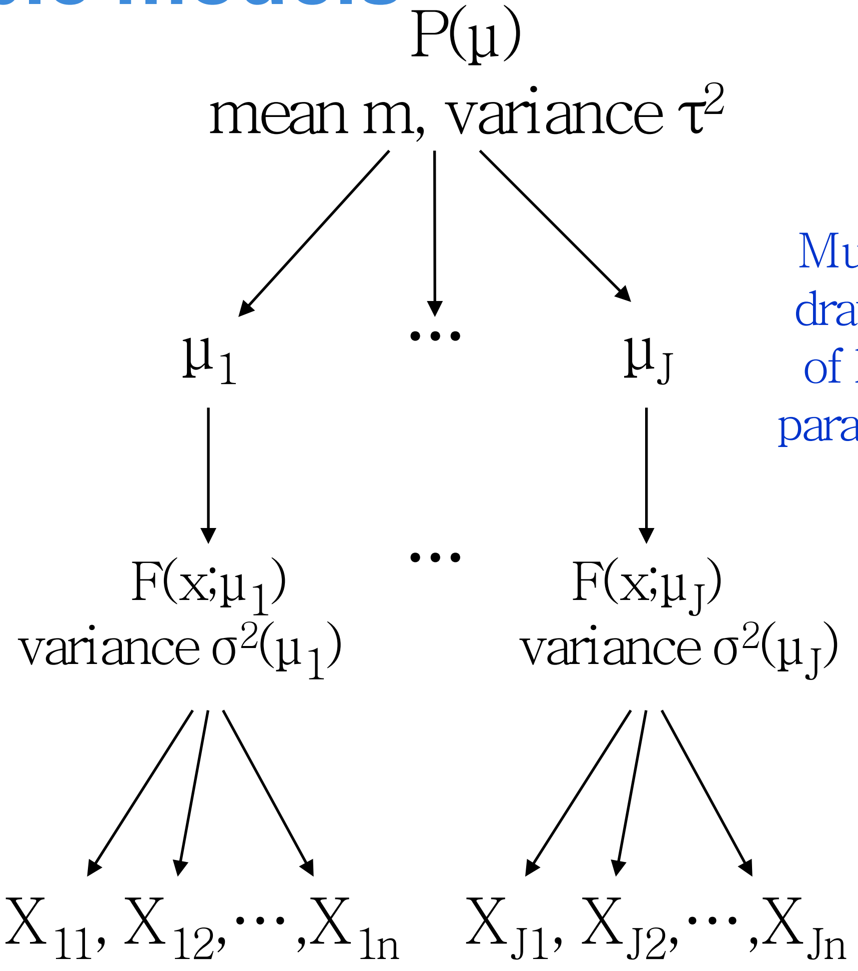
$X_{11}, X_{12}, \cdots, X_{1n}$ $X_{J1}, X_{J2}, \cdots, X_{Jn}$

Multiple drawings of latent parameters

Estimation of credibility coefficients in simple models

Data set-up now

		Observations			
Risk classes	**1**	X_{11}	X_{12}	…	X_{1n}
	2	X_{21}	X_{22}	…	X_{2n}
	J	X_{J1}	X_{J2}	…	X_{Jn}



$P(\mu)$

mean m, variance τ^2

μ_1 … μ_J

$F(x;\mu_1)$ variance $\sigma^2(\mu_1)$ … $F(x;\mu_J)$ variance $\sigma^2(\mu_J)$

$X_{11}, X_{12}, \cdots, X_{1n}$ $X_{J1}, X_{J2}, \cdots, X_{Jn}$

Multiple drawings of latent parameters

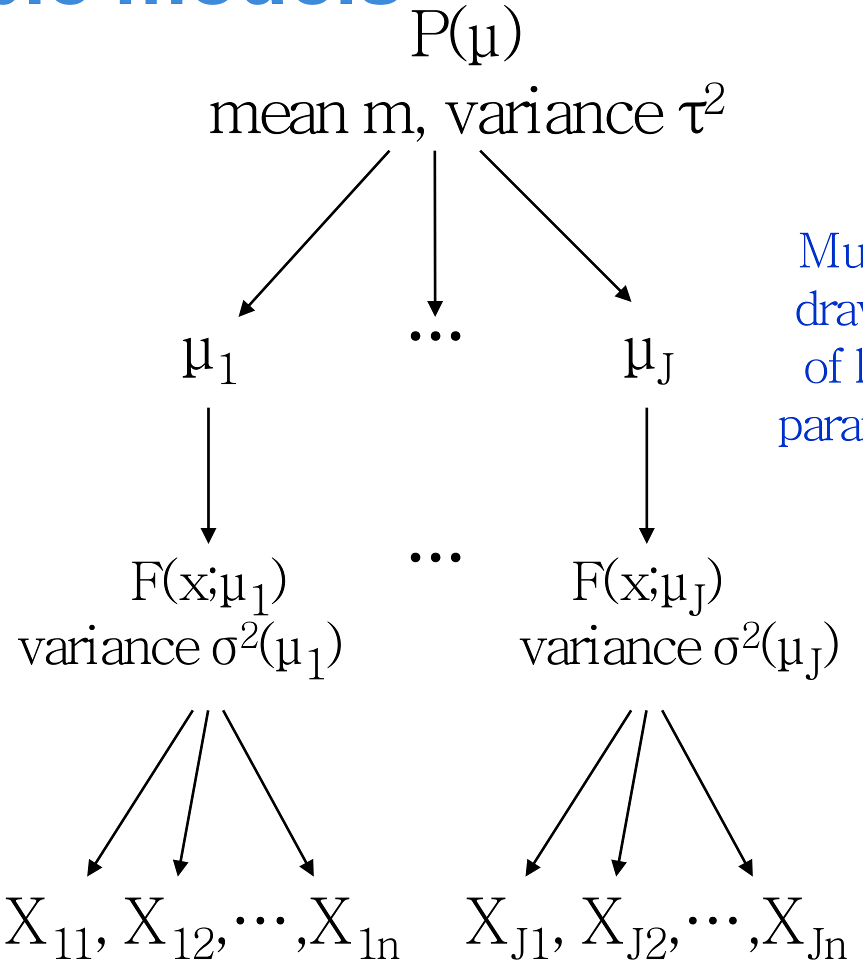
Estimation of credibility coefficients in simple models

Data set-up now

		Observations			
Risk classes	**1**	X_{11}	X_{12}	…	X_{1n}
	2	X_{21}	X_{22}	…	X_{2n}
	:				
	J	X_{J1}	X_{J2}	…	X_{Jn}

Required quantity $E_\mu[\sigma^2(\mu)]/\tau^2$ is

$$\frac{\text{Within-class variance}}{\text{Between-class variance}}$$



$P(\mu)$

mean m, variance τ^2

μ_1 … μ_J

$F(x;\mu_1)$

variance $\sigma^2(\mu_1)$

…

$F(x;\mu_J)$

variance $\sigma^2(\mu_J)$

$X_{11}, X_{12}, \cdots, X_{1n}$ $X_{J1}, X_{J2}, \cdots, X_{Jn}$

Multiple drawings of latent parameters

Analysis of variance

Special case $\sigma^2(\mu) = \sigma^2$

		Observations			
Risk classes	1	X_{11}	X_{12}	…	X_{1n}
	2	X_{21}	X_{22}	…	X_{2n}
	J	X_{J1}	X_{J2}	…	X_{Jn}

Required quantity $E_\mu[\sigma^2(\mu)]/\tau^2 = \sigma^2/\tau^2$ is

$$\frac{\text{Within-class variance}}{\text{Between-class variance}}$$

- This requires an **analysis of variance**
- Required ratio estimated by 1/F where F is ANOVA test statistic for null hypothesis

$$H_0: \mu_1 = \ldots = \mu_J$$

- This yields following estimator of credibility coefficient

$$\mathbf{z = (1 + 1/nF)^{-1}}$$

- Proved by Zehnwirth (1977)

Analysis of variance as regression

		Observations			
Risk classes	1	X_{11}	X_{12}	…	X_{1n}
	2	X_{21}	X_{22}	…	X_{2n}
	J	X_{J1}	X_{J2}	…	X_{Jn}

- Write

$$X_{ij} = \mu_j + \varepsilon_{ij}$$

with $E[\varepsilon_{ij}] = 0$, $Var[\varepsilon_{ij}] = \sigma^2$

- This is a linear regression model
- Null hypothesis

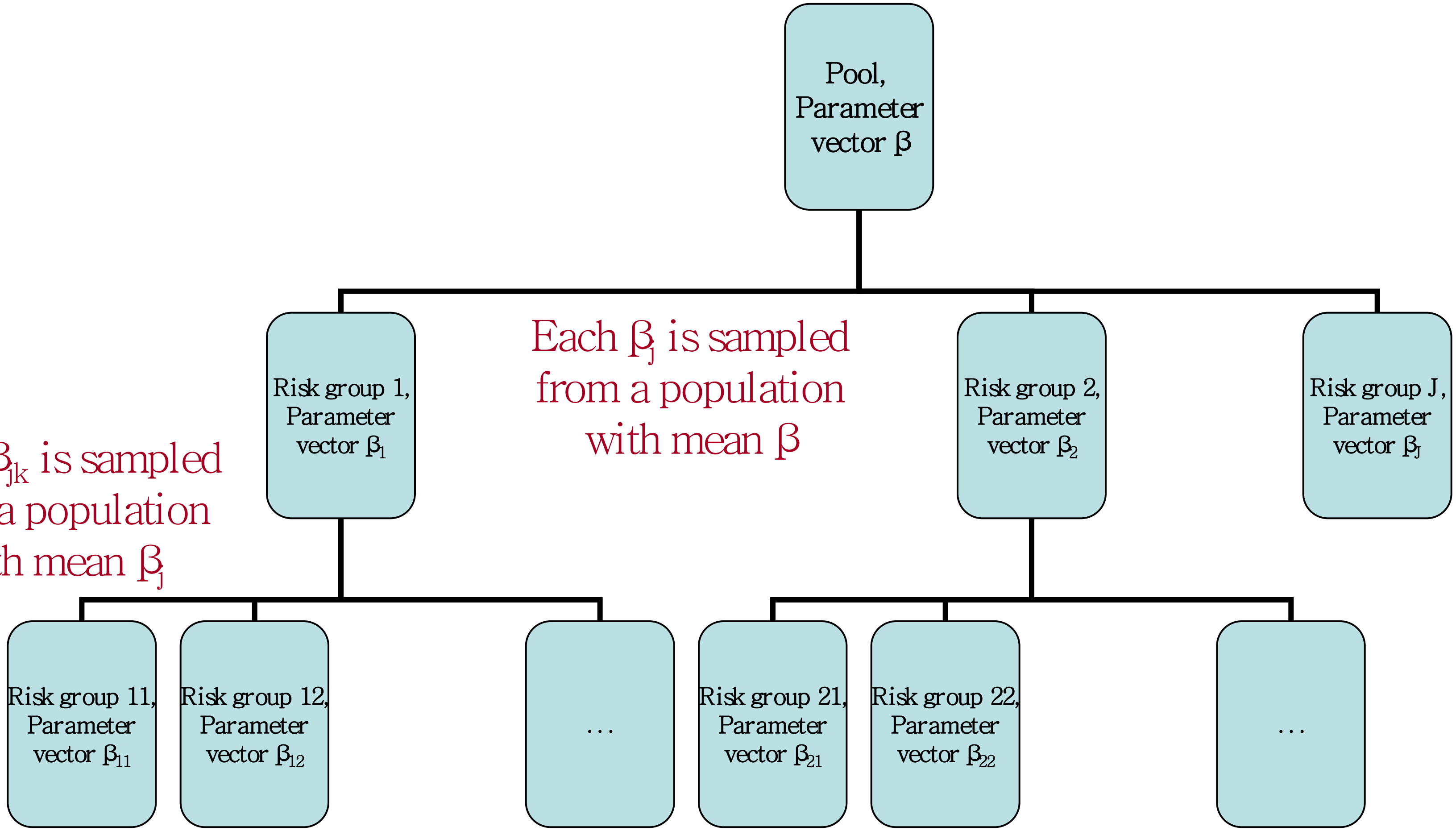
$$H_0: \mu_1 = \ldots = \mu_J$$

- Corresponds to model

$$X_{ij} = \mu + \varepsilon_{ij}$$

- ANOVA F-statistic is same as F-statistic for testing simpler regression model against more complex
  - Credibility coefficient obtainable by performing a regression F-test
  - Regression software can be used

More general model – hierarchical example



Pool, Parameter vector β

Risk group 1, Parameter vector β_1

Risk group 2, Parameter vector β_2

Risk group J, Parameter vector β_J

Each β_j is sampled from a population with mean β

Each β_{jk} is sampled from a population with mean β_j

Risk group 11, Parameter vector β_{11}

Risk group 12, Parameter vector β_{12}

. . .

Risk group 21, Parameter vector β_{21}

Risk group 22, Parameter vector β_{22}

. . .

etc. (eventually observations)

Hierarchical example (cont'd)

- Procedure essentially as for simpler example
- E.g. estimate β_{jk} by

$$\hat{\beta}_{jk} = (1\text{-}z_{jk})\ \hat{\beta}_j + z_{jk}\ \bar{X}_{jk}$$

Pooled mean of all observations that have node (j.k) as root

and

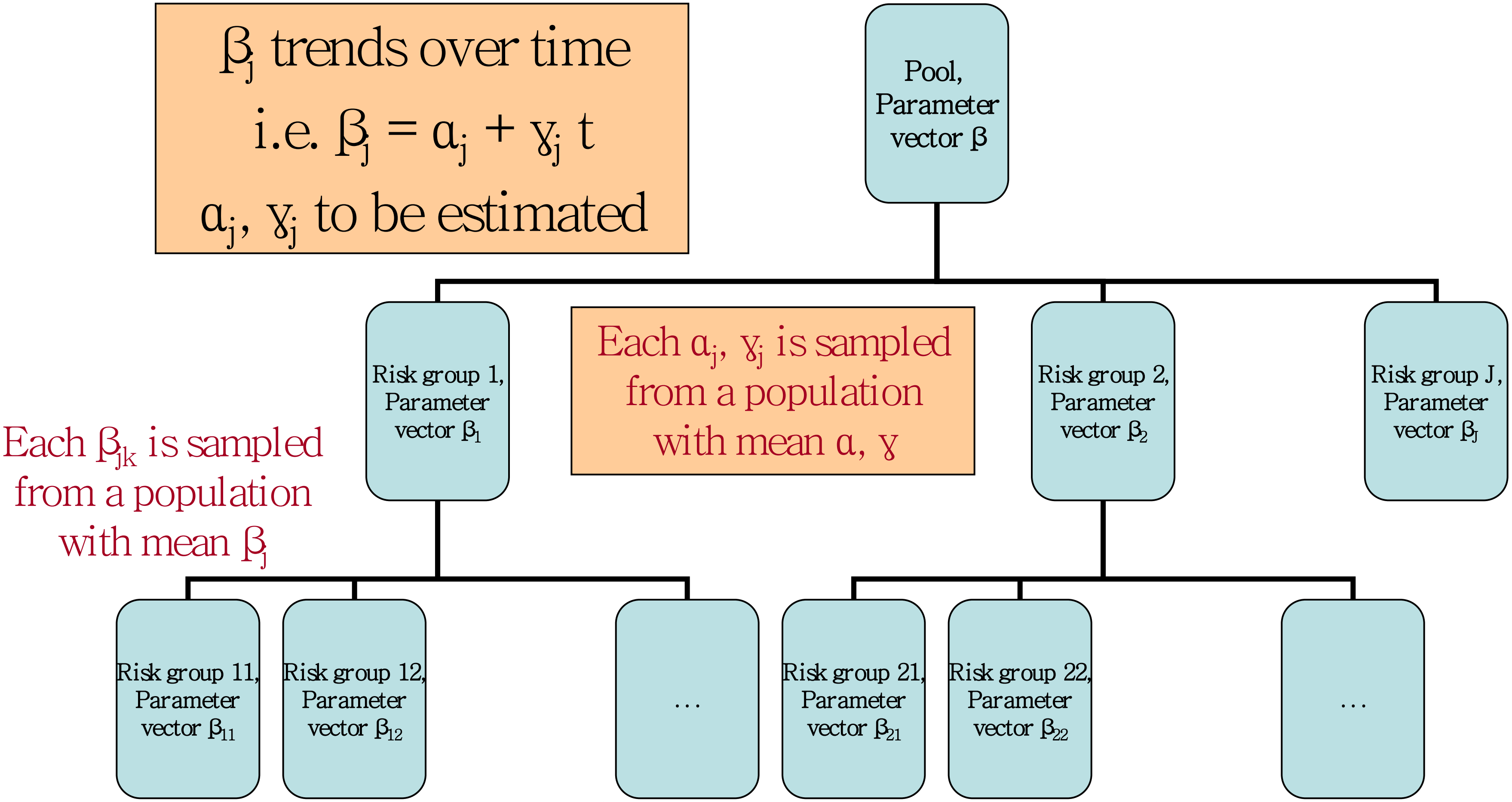
$$z_{jk} = [1 + 1/\ n_{jk}\ F_{jk}]^{-1}$$

Number of observations that have node (j.k) as root

Hierarchical example (cont'd)

- F_{jk} calculated as follows
  - – Define null hypothesis H_0: $\beta_{j1} = \beta_{j2} = \ldots = \beta_{jk} = \ldots$ $(= \beta_j)$
  - – Set up ANOVA with observations on risk classes (j,1), (j,2),…
    - Observations on risk class (j,k) are all those that have node (j,k) as root
  - – F_{jk} is F-statistic for this ANOVA
    - Equivalently regression F-statistic if null hypothesis described in regression terms

Further examples – hierarchical model with trend (Sundt, 1979, 1980)



β_j trends over time

i.e. $\beta_j = \alpha_j + \gamma_j\, t$

α_j, γ_j to be estimated

Pool, Parameter vector β

Risk group 1, Parameter vector β_1

Each α_j, γ_j is sampled from a population with mean α, γ

Risk group 2, Parameter vector β_2

Risk group J, Parameter vector β_J

Each β_{jk} is sampled from a population with mean β_j

Risk group 11, Parameter vector β_{11}

Risk group 12, Parameter vector β_{12}

. . .

Risk group 21, Parameter vector β_{21}

Risk group 22, Parameter vector β_{22}

. . .

etc. (eventually observations)

References (1)

- Gisler A & Müller P (2007). Credibility for additive and multiplicative models. **Astin Colloquium**, Orlando, Fl, USA
- Sundt B (1979). A hierarchical credibility regression model. **Scandinavian Actuarial Journal**, 107-114
- Sundt B (1980). A muti-level hierarchical credibility regression model. **Scandinavian Actuarial Journal**, 25-32

References (2)

- Taylor G C (1979). Credibility analysis of a general hierarchical model. **Scandinavian Actuarial Journal**, 1-12.
- Taylor G (2007). Credibility, hypothesis testing and regression software. **Astin Bulletin**, 37 (in press)
  - Also appears as University of Melbourne Research Paper No. 149 at http://www.economics.unimelb.edu.au/SITE/actwww/wps2007/No149.pdf
- Zehnwirth B (1977). The credible distribution is an admissible Bayes rule. **Scandinavian Actuarial Journal**, 121-127.