(44) <u>Proof</u>: <u>idea</u>: Application of Prop. 4.2.2 (Minimum risk estimation) to the special case of

$X = \mu(\Theta_i) \overset{\text{def}}{=} E[X_{i,1} \mid \Theta_i]$,

$Y = (\underbrace{X_{1,1}, \dots, X_{1,n_1}}_{=X_1}, \dots, \underbrace{X_{r,1}, \dots, X_{r,n_r}}_{=X_r})'$ ← transpose

$a = (\underbrace{a_{1,1}, \dots, a_{1,n_1}}_{=:a_1}, \dots, \underbrace{a_{r,1}, \dots, a_{r,n_r}}_{=:a_r})'$

→ we want to <u>find</u> (a_0, a) w.r.t. the <u>normal equations</u> (4.2.12):

$$a_0 = E[X] - a' \cdot E[Y], \quad \Sigma_{X,Y}' = a' \cdot \Sigma_Y \qquad (*)$$

$$\longrightarrow \text{minimizer } \hat{\mu}_{LB} = \hat{X} = a_0 + a' \cdot Y$$

W.l.o.g. $r = 2$, $n_1 = n_2 = 1$ and $i = 1$

$$\longrightarrow Y = (X_{1,1}, X_{2,1})' \xRightarrow[\text{indep., hence uncorrelated}]{X_{1,1}, X_{2,1}} \Sigma_Y = \begin{pmatrix} Var[X_{1,1}] & 0 \\ 0 & Var[X_{2,1}] \end{pmatrix}$$

and $\Sigma_{X,Y} \overset{\text{def.}}{=} ((Cov[X, X_{1,1}], Cov[\overset{E[X|\Theta_i]}{\overset{\|}{X}}, X_{2,1}]))' = ((Cov[X, X_{1,1}], 0))'$ (indep, $= 0$)

$$\Longrightarrow \text{2nd eq. of } (*) \Longleftrightarrow \overset{\lambda \leftarrow L.4.2.7}{Cov[\mu(\Theta_1), X_{1,1}]} = a_{1,1} \overbrace{Var[X_{1,1}]}^{= \varsigma + \lambda \leftarrow L.4.2.7 (iii)} \qquad (**)$$

and $0 = a_{2,1} \underbrace{Var[X_{2,1}]}_{\overset{L.4.2.7(iii)}{=} \varsigma + \lambda > 0} \Longrightarrow a_{2,1} = 0$

$$E[X_{i,1}] = E[\mu(\Theta_i)] \overset{\text{def}}{=} \mu, \quad \underset{(*)}{\Longrightarrow} \quad a_0 = \mu - (a_{1,1}\mu + a_{2,1}\mu) = \mu(1 - \overbrace{(a_{1,1} + a_{2,1})}^{= \sum_{t=1}^{n_i} a_{i,t}}) \qquad (***)$$

$$\longrightarrow (**) \Longleftrightarrow \lambda = a_{1,1}\varsigma + (a_{1,1} + \underbrace{a_{2,1}}_{=0})\lambda$$

$$\Longrightarrow a_{1,1} = \frac{\lambda}{\varsigma + \lambda} = \frac{\lambda}{\varsigma + \underbrace{n_i}_{=1} \lambda} \qquad \overline{X}_i \overset{\text{def}}{=} \frac{1}{n_i} \sum_{t=1}^{\overset{=1}{n_i}} X_{i,t}$$

$$\Longrightarrow a_0 = \mu \cdot \frac{\varsigma}{\varsigma + n_i \lambda}$$

$$\xrightarrow{W := \overset{=1}{n_i} \cdot a_{1,1}} \hat{\mu}_{LB} = \overset{a_0}{(1-w)\mu} + \overset{w}{a_{1,1}} \cdot \overset{\|}{X_{1,1}} = (1-w)\mu + w \cdot \overline{X}_i \qquad (+)$$

<u>risk $\varsigma(\hat{\mu}_{LB})$</u>: $\varsigma(\hat{\mu}_{LB}) \overset{(4.2.14)}{=} \overbrace{Var[\mu(\Theta_i)]}^{=\lambda} - Var[\hat{\mu}_{LB}]$

$$Var[\hat{\mu}_{LB}] \overset{(+)}{=} Var[w\overline{X}_i] = w^2(E[Var[\overline{X}_i \mid \Theta_i]] + Var[E[\overline{X}_i \mid \Theta_i]])$$

$$\overset{\text{def.}}{=} w^2[\varsigma + \lambda] = \lambda \frac{\overset{=1}{n_i} \lambda}{\varsigma + n_i \lambda}$$

$$\Longrightarrow \varsigma(\hat{\mu}_{LB}) = \lambda - \lambda \frac{n_i \lambda}{\varsigma + n_i \lambda} = (1-w)\lambda \qquad \Longrightarrow \text{proof.}$$

15

<u>Ex. 4.2.10</u> : Consider the heterogeneity model (Def 4.1.2).

Assume $X_{i,t} \sim Pois(\Theta_i)$, $\Theta_i \sim \Gamma(\gamma, \beta)$

$\nwarrow$ number of claims in year i

$\Longrightarrow$ $E[X_{i,t} | \Theta_i] = \Theta_i$ and $Var[X_{i,t}|\Theta_i] = \Theta_i$

and

$E[\Theta_i] = \gamma/\beta$ and $Var[\Theta_i] = \gamma/\beta^2$

$\Longrightarrow$ $\varphi \stackrel{def}{=} E[\underbrace{Var[X_{i,t}|\Theta_i]}_{=\Theta_i}] = \gamma/\beta$,

$\lambda \stackrel{def}{=} Var[\underbrace{\mu(\Theta_i)}_{=\Theta_i}] = \gamma/\beta^2$

$\mu \stackrel{def}{=} E[\underbrace{\mu(\Theta_i)}_{=\Theta_i}] = \gamma/\beta$

$\xrightarrow{\text{Th. 4.2.9}}$ $\hat{\mu}_{LB} = (1-w)\mu + w\,\overline{X}_i$ with

$w = \frac{n_i \lambda}{\varphi + n_i \lambda} = \frac{n_i \gamma/\beta^2}{\gamma/\beta + n_i \gamma/\beta^2} = \frac{n_i}{\beta + n_i}$

$\xrightarrow{\text{Prob.1, Exerc.4}}$ $\hat{\mu}_{LB} = \hat{\mu}_B$

<u>Rem. 4.2.11</u>

(i) $\hat{\mu}_{LB} \neq \hat{\mu}_B$ in general (see e.g. Prob.3, Exerc. 4)

(ii) $w \stackrel{def}{=} \frac{n_i \lambda}{\varphi + n_i \lambda} = \frac{n_i}{\varphi/\lambda + n_i}$ is called <u>credibility weight</u>

$\longrightarrow$ the larger w is the less important is the overall information of the portfolio w.r.t. μ in

$\hat{\mu}_{LB} = (1-w)\mu + w \cdot \overline{X}_i$

$\longrightarrow$ the larger w the more important or <u>credible</u> is the information w.r.t. to data X_i

$\longrightarrow$ <u>1. case</u>: w large, if n_i large

<u>2. case</u>: w large, if φ/λ small

<u>Next</u> we want to discuss a <u>refinement of the model of Bühlmann</u>, that is the <u>Bühlmann-Straub model</u>.

$\longrightarrow$ Bühlmann's model with an extra parameter

$P_{i,t} > 0$,

called <u>risk unit per time unit</u> (e.g. year)

$\longrightarrow$ weights which stand for the knowledge about the <u>volume</u> of $X_{i,t}$

→ e.g. $p_{i,t}$ may correspond to the size of a particular house in connection with fire insurance

or one may interprete $p_{i,t}$ as a risk adjustment parameter w.r.t. subclasses (i.e. accounts) of tariff classes, that is categories of policies for the same type of risk (i.e. homogeneous risk categories)

→ e.g. Swiss Motor Liability Tariff:
1. private cars (1.1 private cars 4-wheeled, 1.2 private cars 3-wheeled, ...) ← further subdivided by e.g. cylinder capacity
2. automobiles for goods transport (2.1 work transports, 2.2 transport business, ...)
3. motor-cycles (3.1 small motor cycles, ...)
4. buses
5. Special risks
6. Short term risks

→ Def. 4.2.12 (Bühlmann-Straub model)
The Bühlmann-Straub model is the Bühlmann model (Def. 4.2.5) where condition (V) is replaced by
(V)' There exist functions μ, v s.t. for all $t \geq 1$:
$$\mu(\Theta_i) = E[X_{i,t} \mid \Theta_i] \text{ and } Var[X_{i,t} \mid \Theta_i] = \frac{v(\Theta_i)}{p_{i,t}},$$
where the weights are pre-specified deterministic risk units.
($\underset{\text{i.i.d.}}{\xrightarrow{\Theta_i, i\geq 1}}$ $\mu(\Theta_i), i \geq 1$ and $v(\Theta_i)$ $i \geq 1$ i.i.d.)

→ Th. 4.2.13 ($\hat{\mu}_{LB}$ w.r.t. the Bühlmann-Straub model)
Set $\mu = E[\mu(\Theta_i)]$, $\lambda = Var[\mu(\Theta_i)]$, $\varphi = E[v(\Theta_i)]$. Suppose that $Var[X_{i,t}] < \infty$, $i,t \geq 1$ and that $\Sigma_{X_i}^{-1}$ exists for all i. Then the linear Bayes estimator $\hat{\mu}_{LB}$ of $\mu(\Theta_i)$ (w.r.t. $\mathcal{L}$ in (4.2.5)) is unique and given by
$$\hat{\mu}_{LB} = (1-w)\mu + w\overline{X}_i,$$
where
$$w = \frac{\lambda \cdot p_i}{\varphi + \lambda \cdot p_i}, \quad \overline{X}_i = \frac{1}{p_i}\sum_{t=1}^{n_i} p_{i,t} X_{i,t} \text{ and } p_i = \sum_{t=1}^{n_i} p_{i,t}$$

47 The corresponding risk is

$$S(\hat{\mu}_{LB}) = (1-w)\lambda$$

Proof: analogous to Th. 4.2.9.

5. Point processes and their applications

The concept of a point process is a generalization of that of e.g. a Poisson process or a cluster point process, which can be e.g. used for the prediction of claim numbers or total claim amounts given the claim number or total claim amount history

In order to introduce this concept, we need some notions:

Recall

Def. 5.1 (measure)

Let $\mathcal{A}$ be a σ-algebra on Ω. Then a function $m : \mathcal{A} \longrightarrow [0,\infty]$ is called measure on $\mathcal{A}$, if

(i) $m(\emptyset) = 0$,

(ii) $A_j \in \mathcal{A}, j \geq 1, A_i \cap A_j = \emptyset, i \neq j \Longrightarrow$
$$m\left(\bigcup_{j\geq 1} A_j\right) = \sum_{j \geq 1} m(A_j) \quad (\sigma\text{-additivity})$$

Ex 5.2: $m = P$ prob. meas. on $\mathcal{A}$

Ex 5.3: $m = \lambda$ Lebesgue-Borel meas. on $\mathcal{A} = \mathcal{B}([0,\infty))$ defined as $\mathcal{B}(\mathbb{R})$

$\longrightarrow$ unique measure s.t. $\lambda([a,b)) = b-a$ for all $a \leq b$

Ex. 5.4 (product measure): $m = m_1 \otimes m_2$, m_i measure on $\mathcal{A}_i$, $\mathcal{A} := \mathcal{A}_1 \otimes \mathcal{A}_2 :=$ smallest σ-alg. containing all sets $A_1 \times A_2$, $A_i \in \mathcal{A}_i$, $i=1,2$, $\Omega := \Omega_1 \times \Omega_2$

$\longrightarrow$ $m = m_1 \otimes m_2$ unique measure on $\mathcal{A}$ s.t.
$$m(A_1 \times A_2) = m_1(A_1)\, m_2(A_2) \quad \text{for all } A_i \in \mathcal{A}_i, i=1,2$$

Def. 5.5 (measurable function)

A function $f : \Omega \longrightarrow \mathbb{R}$ is called $\mathcal{A}$-measurable, if
$$\{\omega \in \Omega : f(\omega) < t\} \in \mathcal{A}$$
for all $t \in \mathbb{R}$.

Ex: $f = X$ random variable w.r.t. $\mathcal{A}$

Def. 5.6 (integral w.r.t. m)

The integral of a $\mathcal{A}$-measurable function f w.r.t. the m, denoted by $\int_\Omega f(\omega)\, m(d\omega)$ can be defined