Table

	a)	b)	c)	d)	e)	f)	g)	h)	„Σ"
net premium pr.	+	+	+	+	+	+	+	+	8
expect. pr.	−	+	−	−	+	+	+	−	4
variance pr.	−	+	+	+	−	+	−	−	4
standard dev. pr.	−	+	+	+	+	−	+	−	5
quantil pr.	+	−	+	+	+	−	−	−	4
exponential pr.	+	+	+	+	−	+	+	+	7

- only the net pr. principle has 8 + : but a.s. ruin!
- very good: exponential principle (but does not apply to heavy tailed case)

Basics of Credibility Theory

Ex: $N = 10$ risks over 5 periods

X_{ij} ... risk i in the period j:

X_{ij}	1	2	3	4	5
1	0	0	0	0	1
2	0	0	0	0	0
3	0	1	0	1	0
4	0	0	1	0	0
5	0	1	0	0	1
6	0	0	1	0	0
7	0	0	0	0	0
8	0	0	0	1	0
9	0	1	1	0	1
10	0	0	0	0	0

assume: a unique premium of 0.22 / risk and period ⇒

		premium	benefit	
⇒ after 5 years:	risk 9:	1.1	3	☺
	risk 10:	1.1	0	☹

⇒ one needs to adjust the premiums to the claim history

Suggestion 1: Use only individual data

i.e.: $P_i^I = \frac{1}{n} \sum_{j=1}^{n} X_{ij}$

„each risk is self-sustaining"

- no risk pooling
- premium for the first periods ???
- complete heterogeneous portfolio

Suggestion 2: use only the collective claim history

i.e.: $$p_i^K = \frac{1}{n \cdot N} \cdot \sum_{i=1}^{N} \sum_{j=1}^{n} X_{ij}$$

- assumes a very homogeneous portfolio: unrealistic
- „not fair", see above

Rem: there are more collective data, but the individual ones are more relevant!

Basic idea: Combine the two suggestions by a <u>weighted average</u>

$$P_i := z \cdot P_i^I + (1-z) P_i^K , \qquad z \in [0,1]$$

Credibility formula

z ... credibility factor : weight of the individual claim history!

Questions:

- estimator for p^I, p^K
- calculation of the cred. factor with the following properties:
  - the more individual data, the higher is z!

- z should depend only on the number of individual data, but <u>not</u> on the data itself

- the more relevant are the collective data, the smaller is z

- possible disadvantage: ins. company has ~~to~~ a greater input to calculate the premiums!

<u>further problems:</u>

- first premium ?

  → classification of risks in subclasses w.r.t. certain characteristics (e.g. type of car, PS number, age of the driver, ...)

  → important: choose the right characteristics by statistical methods

- optimal would be: complete homogeneous subclasses

- a posteriori classification: inclusion of data, which are not known a priori (km/year, ...)

Exact Credibility Theory

Ex: consider a risk over a couple of periods with unknown, time-indep., Poisson parameter λ

i.e.: $X_j \sim Po(\lambda)$, $j = 1, 2, \ldots$

for $j = 1, \ldots, n$ the realisations x are known!

- additional knowledge: $\lambda \in [60, 140]$ with prob. 80 %

  a-priori knowledge from collective data!

Wanted: premium for X_{n+1}, i.e. estimator for $\mathbb{E}[X_{n+1}]$??

should include:
- a priori information
- a posteriori information $\underline{x}$

- often: apriori distribution of λ: Γ distribution
  - conjugated to the Poisson distr.

- Choose Γ-distr. Λ, s.t.: $\mathbb{P}(60 \leq \Lambda \leq 140) = 0.8$ (not unique)

one has:

$$\left.\begin{aligned} \pi(\lambda) &= \frac{\beta^\alpha}{\Gamma(\alpha)} \cdot \lambda^{\alpha-1} \cdot e^{-\lambda\beta} \\ \mathbb{P}(X=x \mid \Lambda=\lambda) &= e^{-\lambda} \cdot \frac{\lambda^x}{x!} \end{aligned}\right\} \Longrightarrow \text{see chapter about Bayes statistics}$$

$$\pi(\lambda \mid \underline{x}) \overset{prop}{\sim} \lambda^{\alpha + \sum_{i=1}^{n} x_i - 1} \cdot e^{-\lambda(\beta+n)} \quad \hat{=} \quad \Gamma(\alpha + n\bar{x}, \beta + n)$$

Premium: <u>conditional expectation</u> of $X/_\Lambda$ <u>averaged</u> with <u>the aposteriori distribution of Λ</u>

$$p = \mathbb{E}[\,\mathbb{E}[X \mid \Lambda] \mid \underline{X}] \underset{Ex.}{=} \frac{\alpha + n\bar{x}}{\beta + n}$$

compare: p^k ... <u>conditional $\mathbb{E}$</u> of $X/_\Lambda$ <u>averaged with</u> <u>a-priori distribution</u>: $\mathbb{E}[\,\mathbb{E}[X \mid \Lambda]] \underset{ex}{=} \dfrac{\alpha}{\beta}$

p^I ... $\bar{X}$

$$\Rightarrow \quad P = \frac{\alpha + n\bar{x}}{\beta + n} = \frac{\alpha}{\beta} \cdot \frac{\beta}{\beta + n} + \bar{x} \cdot \frac{n}{\beta + n}$$

<u>Credibility form !!</u> with $z = \frac{n}{n+\beta}$

<u>One has</u>:
- affine w.r.t. $\bar{x}$
- $z \in [0,1]$ $\quad (\beta > 0 !)$
- $n \to \infty \Rightarrow z \to 1$
- $\beta \to \infty$: a priori distribution: $V(\Lambda) = \frac{\alpha}{\beta^2} \longrightarrow 0$

  $\Rightarrow$ relevance of the collective data increases,

  hence: $z \to 0$ ✓
- z depends on the data only on their number n, not on the data itself.

<u>in general:</u> $X|_{\Theta=\theta} \triangleq F(x|\theta)$, $\theta \in \mathbb{R}^m$, $\Theta \triangleq \pi(\theta)$

premium: <u>exact credibility estimator</u>

$$\bar{e}(\underline{x}) := \mathbb{E}[\,\mathbb{E}[X|\Theta]\,|\,\underline{X}]$$

<u>procedure</u> either – averaging the $\mathbb{E}[X|\Theta]$ with the a posteriori distribution (see above)

<u>or</u> – <u>expectation w.r.t. predictive distribution</u>, because:

$$\underline{\int_{\mathbb{R}^m} d\theta\, \mathbb{E}[X|\theta]\cdot\pi(\theta|\underline{x})} = \int_{\mathbb{R}^m} d\theta \int_0^\infty dx\; x\cdot f(x|\theta)\cdot\pi(\theta|\underline{x}) = \int_0^\infty dx\; x \int_{\mathbb{R}^m} d\theta\; f(x|\theta)\cdot\pi(\theta|\underline{x}) =$$

$$\underline{\int_0^\infty dx\; x\cdot f(x|\underline{x})}$$

Claim: $\underline{\bar{e}(\underline{x})}$ is the <u>optimal estimator</u> for $\mathbb{E}[X|\theta]$ w.r.t. the <u>quadratic loss function.</u>

Lemma: The best approximation of the r.v. T by a function of $\underline{X}$ in the $L^2(\mathbb{P})$ norm is given by: $\mathbb{E}[T|\underline{X}]$, i.e.

$$\mathbb{E}[(T-h(\underline{X}))^2] \text{ is minimal for } h = \mathbb{E}[T|\underline{X}].$$

Proof. $\mathbb{E}\left[(T-h(\underline{X}))^2\right] = \mathbb{E}\left[\left(T-\mathbb{E}[T|\underline{X}]+\mathbb{E}[T|\underline{X}]-h(\underline{X})\right)^2\right] =$

$$\mathbb{E}\left[(T-\mathbb{E}[T|\underline{X}])^2\right] + \mathbb{E}\left[(\mathbb{E}[T|\underline{X}]-h(\underline{X}))^2\right] + \underbrace{2\mathbb{E}\left[(T-\mathbb{E}[T|\underline{X}])(\mathbb{E}[T|\underline{X}]-h(\underline{X}))\right]}_{(*)} =: (+)$$

$$(*) = \mathbb{E}\left[\mathbb{E}\left[(T-\mathbb{E}[T|\underline{X}])(\mathbb{E}[T|\underline{X}]-h(\underline{X}))\,|\,\underline{X}\right]\right] =$$

$$\mathbb{E}\Big[\underbrace{\mathbb{E}\left[(T-\mathbb{E}[T|\underline{X}])\,|\,\underline{X}\right]}_{=0}\cdot(\mathbb{E}[T|\underline{X}]-h(\underline{X}))\Big] = 0$$

$\Rightarrow$ $(+)$ is minimal for $h(\underline{X}) = \mathbb{E}[T|\underline{X}]$ (2nd term $= 0$!)

Application: $T = \mathbb{E}[X|\theta]$

$\Rightarrow h(\underline{X}) = \mathbb{E}[T|\underline{X}] = \mathbb{E}[\mathbb{E}[X|\theta]|\underline{X}]$ ✓

Summary:

- model: $X_i|_{\Theta=\theta} \overset{\wedge}{=} F(x|\theta)$, $\Theta \overset{\wedge}{=} \pi(\theta)$
- $X_i|_{\Theta=\theta}$ are independent (see the form of the likelihood: $L(\theta|\underline{x}) = \prod f(x_i|\theta)$!)
- X_i are not independent!

Example: Poisson-Γ model

$$\mathbb{E}[X_1X_2] = \mathbb{E}[\mathbb{E}[X_1X_2|\Lambda]] \overset{\text{cond. indep.}}{=} \mathbb{E}[\mathbb{E}[X_1|\Lambda]\,\mathbb{E}[X_2|\Lambda]] = \mathbb{E}[\Lambda^2] \overset{\Gamma\text{-distr.}}{=}$$

$$= \frac{\alpha}{\beta^2} + \frac{\alpha^2}{\beta^2}$$

$$\mathbb{E}[X_i] = \mathbb{E}[\mathbb{E}[X_i|\Lambda]] = \mathbb{E}[\Lambda] = \frac{\alpha}{\beta}$$

$$\Rightarrow \operatorname{cov}(X_1, X_2) = \mathbb{E}[X_1X_2] - \mathbb{E}[X_1]\mathbb{E}[X_2] = \frac{\alpha}{\beta^2} + \frac{\alpha^2}{\beta^2} - \left(\frac{\alpha}{\beta}\right)^2 = \frac{\alpha}{\beta^2} \neq 0 \quad \blacksquare$$

Remarks:
- the exact credibility premium has <u>not</u> always credibility form!
- — " — : sometimes difficult to calculate
- Which a priori distribution should one choose ???

Empirical Credibility Theory

Bühlmann model:

Problem: Estimate $\mathbb{E}[X_{n+1}]$ using $\underline{X}$

Assumptions:

- Distribution of $X_1, \dots, X_{n+1}$ depends on a certain parameter θ (risk parameter)
- θ is independent of time but random var.
- $m(\theta) := \mathbb{E}[X_j \mid \theta]$ $\quad v(\theta) := \mathbb{V}(X_j \mid \theta)$

  indep. of time, hence indep. of j
- $X_1, \dots, X_{n+1}$ are conditional uncorrelated: $\mathbb{E}[X_i X_j \mid \theta] = \mathbb{E}[X_i \mid \theta]\, \mathbb{E}[X_j \mid \theta]$

  $i \neq j$

Differences to exact credibility theory:

- no assumptions on the distribution of $X_i | \theta$ and θ
- conditional uncorrelated instead of cond. indep.: X_i, X_j

– if θ would be known, one would take $m(\theta)$!

– exact cred. theory: $\tilde{e}$ is the estimator for $m(\theta)$, which minimizes the L^2 difference ~~under~~ among all $h(\underline{X})$!

– now: look for the minimizer ~~under~~ among all affine estimators i.e.: $e(\underline{x}) = \alpha_0 + \sum_{i=1}^{n} \alpha_i X_i$

with $\mathcal{E} := \mathbb{E}\left[\left(m(\theta) - \alpha_0 - \sum_{i=1}^{n} \alpha_i X_i\right)^2\right] \longrightarrow \min$

Lemma: One has:

1.) $\mathbb{E}[X_j\, m(\theta)] = \mathbb{V}(m(\theta)) + (\mathbb{E}[m(\theta)])^2$

2.) $\mathbb{E}[X_i X_j] = \quad$ — " — $\quad , \; i \neq j$

3.) $\mathbb{E}[X_i^2] = \mathbb{V}(m(\theta)) + (\mathbb{E}[m(\theta)])^2 + \mathbb{E}[v(\theta)]$

<u>Proof:</u> ad1) $\mathbb{E}[X_j\, m(\theta)] = \mathbb{E}[\mathbb{E}[X_j\, m(\theta)|\theta]] = \mathbb{E}[\mathbb{E}[X_j\, \mathbb{E}[X_j|\theta]|\theta]] =$

$\mathbb{E}[(\underbrace{\mathbb{E}[X_j|\theta]}_{m(\theta)})^2] = \mathbb{V}(m(\theta)) + (\mathbb{E}[m(\theta)])^2$ ✓

ad2) $\mathbb{E}[X_i X_j] = \mathbb{E}[\mathbb{E}[X_i X_j|\theta]] \underset{\uparrow}{=} \mathbb{E}[\mathbb{E}[X_i|\theta]\,\mathbb{E}[X_j|\theta]] = \mathbb{E}[m(\theta)^2]$

cond. uncorrelated

$= \mathbb{V}(m(\theta)) + (\mathbb{E}[m(\theta)])^2$ ✓

ad3) $\mathbb{E}[X_j^2] = \mathbb{E}[\mathbb{E}[X_j^2|\theta]] = \mathbb{E}[(\mathbb{E}[X_j|\theta])^2 + \mathbb{V}(X_j|\theta)] \overset{1.)}{=}$

$\mathbb{V}(m(\theta)) + (\mathbb{E}[m(\theta)])^2 + \mathbb{E}[v(\theta)]$ ✓

□

- <u>Restriction to conditions of first order in the extremal problem above</u>

$$\frac{\partial \mathcal{E}}{\partial \alpha_0} = -2\,\mathbb{E}\Big[m(\theta) - \alpha_0 - \sum_{i=1}^{n} \alpha_i X_i \Big] = 0 \qquad (1)$$

$$\frac{\partial \mathcal{E}}{\partial \alpha_j} = -2\,\mathbb{E}\Big[X_j \Big(m(\theta) - \alpha_0 - \sum_{i=1}^{n} \alpha_i X_i \Big)\Big] = 0 \qquad (2)$$

$$(1) \Rightarrow \underline{\alpha_0} = \mathbb{E}\Big[m(\theta) - \sum_{i=1}^{n} \alpha_i X_i \Big] = \mathbb{E}\Big[\mathbb{E}\big[m(\theta) - \sum \alpha_i X_i \,|\, \theta \big]\Big] =$$

$$\underline{\mathbb{E}\Big[m(\theta) \Big(1 - \sum_{i=1}^{n} \alpha_i \Big)\Big]} \qquad (3)$$

(2) $\rightarrow$

$$\alpha_j \, \mathbb{E}[X_j^2] = \mathbb{E}\Big[X_j \, m(\theta) - \alpha_0 X_j - \sum_{i \neq j} \alpha_i X_i X_j\Big]$$

$$\overset{\text{1.) 2.) 3.) from Lemma above}}{\Longleftrightarrow}$$

$$\alpha_j \Big\{ \mathbb{E}[v(\theta)] + \mathbb{V}(m(\theta)) + (\mathbb{E}[m(\theta)])^2 \Big] = \mathbb{V}(m(\theta)) + (\mathbb{E}[m(\theta)])^2 - \sum_{i \neq j} \alpha_i \Big(\mathbb{V}(m(\theta)) + (\mathbb{E}[m(\theta)])^2 \Big) - \alpha_0 \, \mathbb{E}[m(\theta)]$$

$$\Longleftrightarrow$$

$$\alpha_j \, \mathbb{E}[v(\theta)] = \Big(\mathbb{V}(m(\theta)) + (\mathbb{E}[m(\theta)])^2 \Big) \Big\{ 1 - \sum_{i=1}^{n} \alpha_i \Big\} - \alpha_0 \, \mathbb{E}[m(\theta)] \qquad (4)$$

$\Rightarrow \alpha_j$ indep. of j; let $\alpha_j =: \frac{z}{n} \overset{(3)}{\Longrightarrow}$

$$\underline{\alpha_0 = \mathbb{E}[m(\theta)(1-z)]} \qquad (5)$$

$(4) \rightarrow$

$$\underline{\mathbb{E}[v(\theta)]\cdot\frac{z}{n}} = \left(\mathbb{V}(m(\theta)) + \left(\mathbb{E}[m(\theta)]\right)^2\right)\cdot(1-z) - \alpha_0\cdot\mathbb{E}[m(\theta)] \overset{(5)}{=}$$

$$\left(\mathbb{V}(m(\theta)) + \left(\mathbb{E}[m(\theta)]\right)^2\right)(1-z) - (1-z)\left(\mathbb{E}[m(\theta)]\right)^2 =$$

$$\underline{\mathbb{V}(m(\theta))\,(1-z)} \qquad (6)$$

$$\Rightarrow z\cdot\left(\frac{\mathbb{E}[v(\theta)]}{n} + \mathbb{V}(m(\theta))\right) = \mathbb{V}(m(\theta)) \Rightarrow$$

$$\underline{\underline{z}} = \frac{\mathbb{V}(m(\theta))}{\frac{\mathbb{E}[v(\theta)]}{n} + \mathbb{V}(m(\theta))} = \frac{1}{\frac{\mathbb{E}[v(\theta)]}{n\cdot\mathbb{V}(m(\theta))} + 1} = \underline{\underline{\frac{n}{n + \frac{\mathbb{E}[v(\theta)]}{\mathbb{V}(m(\theta))}}}}$$

$$\Rightarrow$$

Theorem (Bühlmann): Under the assumptions above we have:

The affine estimator for $m(\theta)$, which minimizes the quadratic error is given by:

$$e(\underline{X}) = (1-z)\,\mathbb{E}[m(\theta)] + z \cdot \bar{X} \quad \text{with}$$

$$z = \frac{n}{n + \frac{\mathbb{E}[v(\theta)]}{\mathbb{V}(m(\theta))}}$$

Rem: • Credibility form with $p^I = \bar{X}$, $p^K = \mathbb{E}[m(\theta)]$

- $\lim_{n \to \infty} z = 1$

- $0 \le z \le 1$

- $\mathbb{E}[v(\theta)]$ ... averaged variance of the risks (described by the r.v. $X|_{\Theta=\theta}$)

  : $\mathbb{E}[v(\theta)] \nearrow$ means individual data are more volatile $\Rightarrow z \searrow$ ✓

- $V(m(\theta))$: measure for the deviation of the single risks from each other

  : $V(m(\theta)) \nearrow \Rightarrow$ collective data are less reliable $\Rightarrow z \nearrow \Rightarrow (1-z) \searrow$ ✓

Estimators for $\mathbb{E}[m(\theta)]$, $\mathbb{V}(m(\theta))$, $\mathbb{E}[v(\theta)]$

Consider N risks X_{ij} over n periods $\quad 1 \le i \le N$

$1 \le j \le n$

Assumptions:

- X_{ij} depend on θ_i (realisation of Θ_i), time indep.
- the random vectors $(\Theta_i, X_{i1}, \dots, X_{in})$, $i = 1, \dots, N$ are iid
- $\left.\begin{array}{l} \mathbb{E}[X_{ij} | \theta_i] = m(\theta_i) \\ \mathbb{V}(X_{ij} | \theta_i) = v(\theta_i) \end{array}\right\}$ indep. of i, j, i.e. e.g. $m(\Theta_i)$ is a r.v. with the same distr. $\forall i$
- $\operatorname{cov}(X_{ij}, X_{ik} | \Theta_i) = 0 \quad j \neq k$ ... uncorrelated in time

$\Rightarrow$ for each risk : conditions of Bühlmann Theorem are satisfied

$\Rightarrow$ premium for risk i : $e(X_{i1}, \dots, X_{in}) = (1-z)\,\mathbb{E}[m(\Theta_i)] + z\bar{X}_i$

$$\text{with } z = \frac{n}{n + \frac{\mathbb{E}[v(\Theta_i)]}{\mathbb{V}(m(\Theta_i))}} \text{ , indep. of } i \text{ !}$$

Let now $\bar{X}_i := \frac{1}{n}\sum_{j=1}^{n} X_{ij}$ … average of risk i

- Estimator for $\mathbb{E}[m(\Theta_i)] =: m$

average over all risks: $\bar{X} := \frac{1}{N} \bar{X}_i = \frac{1}{nN} \sum_{i=1}^{N} \sum_{j=1}^{n} X_{ij}$

Claim: $\bar{X}$ is unbiased for m

Proof: to show: $\mathbb{E}[\bar{X}] = m$

$$\mathbb{E}[\bar{X}] = \frac{1}{nN} \sum_{i=1}^{N} \sum_{j=1}^{n} \mathbb{E}[X_{ij}] = \frac{1}{nN} \sum_{i=1}^{N} \sum_{j=1}^{n} \mathbb{E}[\mathbb{E}[X_{ij} | \Theta_i]] =$$

$$\frac{1}{nN} \sum_{i=1}^{N} \sum_{j=1}^{n} \mathbb{E}[m(\Theta_i)] \underset{\text{ass!}}{=} \underset{\text{indep. of } i}{\mathbb{E}[m(\Theta_i)]} = m \quad \checkmark$$

- Estimator for $\mathbb{E}[v(\Theta_i)] =: v$

average over the single sample variances

$$V := \frac{1}{N} \sum_{i=1}^{N} \frac{1}{(n-1)} \sum_{j=1}^{n} (X_{ij} - \bar{X}_i)^2$$

Claim: V is unbiased for v

Proof: to show: $\mathbb{E}[V] = v$

$$\mathbb{E}[V] = \frac{1}{N} \sum_{i=1}^{N} \mathbb{E}\Big[\underbrace{\mathbb{E}\Big[\frac{1}{(n-1)} \sum_{j=1}^{n} (X_{ij} - \bar{X}_i)^2 \,\Big|\, \Theta_i \Big]}_{= v(\Theta_i)} \Big] =$$

(X_{ij} are conditional uncorrelated with variance $v(\Theta_i)$!)

$$= \frac{1}{N} \sum_{i=1}^{N} \mathbb{E}[v(\Theta_i)] = \Big| \, v(\Theta_i) \text{ indep. of } i \, \Big| = \mathbb{E}[v(\Theta_i)] = v \; \checkmark$$

- estimator for $\mathbb{V}(m(\Theta_i)) =: w$

- start from $\frac{1}{N-1}\sum_{i=1}^{N}(\bar{X}_i - \bar{X})^2$ ... is <u>biased</u>, since

- $\underline{\mathbb{E}[\bar{X}_i]} = \mathbb{E}[\mathbb{E}[\bar{X}_i|\Theta_i]] = \frac{1}{n}\mathbb{E}[\sum_{j=1}^{n}\mathbb{E}[X_{ij}|\Theta_i]] = \frac{1}{n}\mathbb{E}[\sum_{j=1}^{n} m(\Theta_i)] = \mathbb{E}[m(\Theta_i)] = \underline{m}$

- $\underline{\mathbb{V}(\bar{X}_i)} = \mathbb{V}(\mathbb{E}[\bar{X}_i|\Theta_i]) + \mathbb{E}[\mathbb{V}(\bar{X}_i|\Theta_i)] =$

$$\mathbb{V}(m(\Theta_i)) + \mathbb{E}[\mathbb{V}(\frac{X_{i1}+\dots+X_{in}}{n}|\Theta_i)] =$$

$$\mathbb{V}(m(\Theta_i)) + \mathbb{E}[\frac{v(\Theta_i)}{n}] = \underline{w + \frac{v}{n}} \quad \Rightarrow$$

$\Rightarrow \bar{X}_1, \ldots, \bar{X}_N$ are indep. with the same mean m and the same variance $w + \frac{v}{n}$ $\Rightarrow$

Unbiased estimator for $w + \frac{v}{n}$: $\frac{1}{(N-1)} \sum_{i=1}^{N} (\bar{X}_i - \bar{X})^2 \Rightarrow$

<u>Unbiased estimator for w:</u> $\underline{W := \frac{1}{N-1} \sum_{i=1}^{N} (\bar{X}_i - \bar{X})^2 - \frac{1}{n} V}$

<u>Remarks:</u>

- it is possible that W is negative!

  $\Rightarrow$ pragmatic solution: set the estimator $= 0$, i.e. $= \max(W, 0)$

  (is biased!)

- z depends, via the estimators, on the data!

- averaged Credibility premium = collective premium, because:

$$\frac{1}{N}\sum_{i=1}^{N} e(X_{i1}, \dots, X_{in}) = \frac{1}{N}\sum_{i=1}^{N}\Big((1-z)\overbrace{\mathbb{E}[m(\Theta_i)]}^{m} + z\bar{X}_i\Big) = \Big| \, m \longrightarrow \text{estimator } \bar{X} \, \Big|$$

$$= \frac{1}{N}\sum_{i=1}^{N}\big((1-z)\bar{X} + z\bar{X}_i\big) = (1-z)\bar{X} + z\bar{X} = \bar{X} \checkmark$$

Practical procedure (Write X_{ij} as matrix)

1.) for each row: calculate $\bar{X}_i$ and $\frac{1}{(n-1)} \sum_{j=1}^{n} (X_{ij} - \bar{X}_i)^2$

2.) calculate $\bar{X}$ and V by averaging over the rows

3.) $W =$ sample variance $(\bar{X}_i) - \frac{1}{n} V$

4.) Calculate z, using $\bar{X}, V, W$!

5.) Calculate the premium using the theorem of Bühlmann

Heavy tailed distributions

Distribution where the tail is of regular variation

Def: Let $L(x)$ be a pos. measurable function on $(0, \infty)$ and we have

$$\lim_{x \to \infty} \frac{L(cx)}{L(x)} = 1 \quad \forall c > 0 \Rightarrow$$

L is called slowly varying (i.e $L \in \mathcal{L}$)

Ex: • each function converging to a positive limit!

Proof: Let $L(x) \xrightarrow{x \to \infty} \alpha > 0 \Rightarrow$

$$\lim_{x \to \infty} \frac{L(cx)}{L(x)} = \frac{\lim_{x \to \infty} L(cx)}{\lim_{x \to \infty} L(x)} = \frac{\alpha}{\alpha} = 1 \checkmark$$

Remark: wrong, if $\alpha = 0$:

e.g.: $h(x) = \frac{1}{x}$: $\lim_{x\to\infty} \frac{h(cx)}{h(x)} = \lim_{x\to\infty} \frac{\frac{1}{cx}}{\frac{1}{x}} = \frac{1}{c} \neq 1$ in general

- $L(x) = \ln x$

Proof: $\lim_{x\to\infty} \frac{\ln(cx)}{\ln(x)} = \lim_{x\to\infty} \frac{\ln c + \ln x}{\ln x} = \lim_{x\to\infty} \left(\frac{\ln c}{\ln x} + 1\right) = 1$ ✓

- similarly: all iterated logaritms and powers of logs

Rem: $\exists\, L \in \mathcal{L}$, s.t.: $\liminf_{x\to\infty} L(x) = 0$, $\limsup_{x\to\infty} L(x) = \infty$

works, as long as the oscillation is slow enough! (e.g: a proper combination of sin and log)

Def: if $\bar{F}(x) = \frac{L(x)}{x^\alpha} = 1 - F(x)$, with $L \in \mathcal{L}$, $\alpha > 0$

$\Rightarrow$ $\bar{F}$ is of regular variation with index α, i.e $\bar{F} \in \mathcal{R}_\alpha$

Examples:

$$\left.\begin{array}{ll} \bullet \text{ Pareto: } & \bar{F}(x) = \left(\frac{k}{k+x}\right)^\alpha, \quad \alpha, k > 0 \\ \\ \bullet \text{ Burr: } & \bar{F}(x) = \left(\frac{k}{k+x^\tau}\right)^\alpha, \quad \alpha, k, \tau > 0 \\ \\ \bullet \text{ Loggamma: } & f(x) = \frac{\alpha^\beta}{\Gamma(\beta)} (\ln x)^{\beta-1} x^{-\alpha-1} \end{array}\right\} \text{ see exercises!}$$

Remarks:

- If $f \in R_\alpha \Rightarrow \bar{F} \in R_{\alpha-1}$

  e.g.: $f(x) = c \cdot x^{-2} \ln x,\ x \geq 1 \Rightarrow \bar{F}(x) = \int_x^\infty f(z)\,dz \in R_1$

- reverse direction is false in general

- heuristically: functions in $R \approx$ like powers

Theorem (without proof): Each function $L \in R$ can be written as

$$L(x) = c(x) \cdot \exp\left(\int_{x_0}^{x} \frac{\varepsilon(t)}{t}\,dt \right) \quad \text{for } x \geq x_0,\ x_0 > 0$$

where $\varepsilon(t) \xrightarrow{t \to \infty} 0$, $c_0(x)$ pos., measurable with $c_0(x) \xrightarrow{x \to \infty} \bar{c}_0 > 0$

"representation theorem"

<u>One has:</u> $L(x)$ is „small" in comparison to powers x^{δ}, $\delta > 0$

i.e: $\lim\limits_{x\to\infty} \frac{L(x)}{x^{\delta}} = 0$

<u>Proof:</u> to show: $\lim\limits_{x\to\infty} \left(\ln L(x) - \delta \ln x \right) = -\infty \qquad (*)$

which means, using repr. th.; $\lim\limits_{x\to\infty} \ln c_0(x) + \lim\limits_{x\to\infty} \left(\int_{x_0}^{x} \frac{\varepsilon(t)}{t} dt - \delta \int_1^x \frac{dt}{t} \right) \overset{?}{=} -\infty$

$$\Leftrightarrow \lim_{x\to\infty} \underbrace{\left(\ln c_0(x) + \int_{x_0}^{1} \frac{\varepsilon(t)}{t} dt \right)}_{|\cdot| \le \text{const.}} + \lim_{x\to\infty} \left(\int_1^x \frac{\varepsilon(t) - \delta}{t} dt \right) \overset{?}{=} \infty$$

$\Rightarrow$ sufficient is: $\lim\limits_{x\to\infty} \left(\int_1^x \frac{\varepsilon(t) - \delta}{t} dt \right) = -\infty \qquad (**)$

Let N be large enough, s.t. $|\varepsilon(t)| \le \frac{\delta}{2}$, $t \ge N$!

$$\Rightarrow \int_1^x \frac{\varepsilon(t)-\delta}{t}\,dt = \int_1^N \frac{\varepsilon(t)-\delta}{t}\,dt + \int_N^x \frac{\varepsilon(t)-\delta}{t}\,dt \le$$

$$\le \text{const.}(N) + \int_N^x \left(-\frac{\delta}{2t}\right)dt \xrightarrow{x\to\infty} -\infty \qquad \square$$

Moreover (an analogous proof yields): $X \sim F$, $\bar{F} \in R_\alpha \Rightarrow$

$$\mathbb{E}[X^\delta] = \begin{cases} \infty & \delta > \alpha \\ < \infty & \delta < \alpha \end{cases}$$

for $\delta = \alpha$: no general assertion possible ($L = ???$)

i.e.: for $\bar{F} \in R_\alpha$ not all moments are finite

Consider: sum of n claims:

$$S_n = X_1 + \dots + X_n; X_i \text{ iid} \quad \text{with} \quad \overline{F} \in R_\alpha$$

Question: asymptotic behavior of S_n for $x \to \infty$??

Lemma: assume: X_1, X_2 indep. with distr. functions F_i, $i = 1,2$ and

$$\overline{F}_i(x) = \frac{L_i(x)}{x^\alpha}, \quad L_i \in \mathcal{L} \quad \text{(possibly different } L\text{'s!)}$$

$\Rightarrow X_1 + X_2$ "is" of regular variation with index α, more precisely:

$$\mathbb{P}(X_1 + X_2 > x) = \left(\mathbb{P}(X_1 > x) + \mathbb{P}(X_2 > x)\right) \cdot (1 + o(1))$$

$$= x^{-\alpha}\left(L_1(x) + L_2(x)\right) \cdot (1 + o(1)), \quad x \to \infty$$

Proof: Let $G(x) := \mathbb{P}(X_1 + X_2 \leq x)$

one has: $\{X_1 + X_2 > x\} \supset \{X_1 > x\} \cup \{X_2 > x\}$

x_i indep

$$\Rightarrow \mathbb{P}(\{X_1 + X_2 > x\}) \geq \mathbb{P}(\{X_1 > x\}) + \mathbb{P}(\{X_2 > x\}) - \underbrace{\mathbb{P}(\{X_1 > x\}) \cdot \mathbb{P}(\{X_2 > x\})}_{= o(\bar{F}_1(x)) = o(\bar{F}_2(x))\ !}$$

$$\Rightarrow \bar{G}(x) \geq (\bar{F}_1(x) + \bar{F}_2(x))(1 - o(1)) \qquad (1)$$

Let now: $0 < \delta < \frac{1}{2} \Rightarrow$

$$\{X_1 + X_2 > x\} \subset \{X_1 > (1-\delta)x\} \cup \{X_2 > (1-\delta)x\} \cup \{X_1 > \delta x,\ X_2 > \delta x\}$$

$$\Rightarrow \bar{G}(x) \le \bar{F}_1((1-\delta)x) + \bar{F}_2((1-\delta)x) + \bar{F}_1(\delta x)\bar{F}_2(\delta)$$

$$= [\bar{F}_1((1-\delta)x) + \bar{F}_2((1-\delta)x)] \cdot (1 + o(1)), \text{ because } \frac{\bar{F}_1(\delta x)}{\bar{F}_1((1-\delta)x)} = \frac{L_1(\delta x)\delta^{-\alpha}}{L_1((1-\delta)x)\cdot(1-\delta)^{-\alpha}}$$

$$= o(x^{\frac{\alpha}{2}})\ !$$

$$\Rightarrow \limsup_{x \to \infty} \frac{\bar{G}(x)}{\underbrace{\bar{F}_1((1-\delta)x)}_{\sim (1-\delta)^{-\alpha}\bar{F}_1(x)} + \underbrace{\bar{F}_2((1-\delta)x)}_{\sim (1-\delta)^{-\alpha}\bar{F}_2(x)}} \le 1$$

since: $\bar{F}_1((1-\delta)x) = \cancel{(1-\delta)^{-\alpha}}\,\cancel{x^{-\alpha}}\,L_1((1-\delta)x) \overset{?}{\sim} \cancel{(1-\delta)^{-\alpha}}\,\cancel{x^{-\alpha}}\,L_1(x)$ ✓

$$\Rightarrow \limsup_{x\to\infty} \frac{\bar{G}(x)}{\bar{F}_1(x)+\bar{F}_2(x)} \le (1-\delta)^{-\alpha} \qquad (2)$$

$$(1)+(2) \Rightarrow 1 \le \liminf_{x\to\infty} \frac{\bar{G}(x)}{\bar{F}_1(x)+\bar{F}_2(x)} \le \limsup_{x\to\infty} \frac{\bar{G}(x)}{\bar{F}_1(x)+\bar{F}_2(x)} \le (1-\delta)^{-\alpha}$$

$$\overset{\delta\to 0}{\Rightarrow} \lim_{x\to\infty} \frac{\bar{G}(x)}{\bar{F}_1(x)+\bar{F}_2(x)} = 1 \qquad \square$$

Corollary: $X_1,\dots,X_n$ iid with distr. function F, s.t. $\bar{F} \in R_\alpha$

$$\Rightarrow S_n = X_1 + \dots + X_n \text{ "}\in\text{" } R_\alpha$$

and $$P(S_n > x) = n\bar{F}(x)(1+o(1)), \quad x\to\infty \qquad (*)$$

Let now : $M_n = \max(X_1, \dots, X_n) \Rightarrow$

$$\mathbb{P}(M_n > x) = (1 - \mathbb{P}(M_n \le x)) = 1 - \mathbb{P}(X_i \le x)^n = 1 - F(x)^n =$$

$$1 - (1 - \bar{F}(x))^n = 1 - \left(1 - n \cdot \bar{F}(x) + \binom{n}{2} \bar{F}(x)^2 - \dots \right)$$

$$= n \bar{F}(x) + o(\bar{F}(x)), \qquad x \to \infty \qquad (+)$$

$$(*)(+) \Rightarrow \boxed{\frac{\mathbb{P}(M_n > x)}{P(S_n > x)} \xrightarrow{x \to \infty} 1, \qquad n \ge 2}$$

which means: for distributions in $\mathcal{R}$: the asymptotic behavior of S_n and M_n is the same!

Subexponential Distributions

Take the last result as <u>starting point</u> $\Rightarrow$

Def: The positive r.v. X has a <u>subexponential distribution</u> F, i.e. <u>$F \in S$</u> $(X \in S)$, if for an iid sequence $X_i \sim F$ one has:

$$\mathbb{P}(S_n > x) = \mathbb{P}(M_n > x) \cdot (1 + o(1)) \qquad \forall n \geq 2, \; x \to \infty$$

Remarks: 1.) One can show that it is enough to impose ✓ for $n = 2$.

This implies that it holds also for $n \geq 2$.

2.) One has (see above): $\mathbb{P}(M_n > x) = n\,\bar{F}(x)(1+o(1))$

$\Rightarrow$ an equivalent def. for $F \in \mathcal{S} \iff \lim\limits_{x\to\infty} \frac{\mathbb{P}(S_n > x)}{\bar{F}(x)} = n \quad , \quad \forall n \geq 2$

Some properties of $F \in \mathcal{S}$

Lemma: (1) $F \in \mathcal{S} \Rightarrow \forall y > 0$:

$$\lim_{x\to\infty} \frac{\bar{F}(x-y)}{\bar{F}(x)} = 1 \qquad (*)$$

(2) If $(*)$ holds, one has $\forall \varepsilon > 0$

$$e^{\varepsilon x} \cdot \bar{F}(x) \longrightarrow \infty \, , \quad x \to \infty$$

„ $\bar{F}$ tends slower to 0 as every exp. function"

Proof: ad (1): Let $G(x) := \mathbb{P}(X_1 + X_2 \le x)$

For $x \ge y > 0$ one has:

$$\frac{\bar{G}(x)}{\bar{F}(x)} = \frac{1 - \int_0^x F(x-t)\, dF(t)}{\bar{F}(x)} = \frac{1 - \int_0^x (1 - \bar{F}(x-t))\, dF(t)}{\bar{F}(x)} =$$

$$\frac{1 - F(x) + \int_0^x \bar{F}(x-t)\, dF(t)}{\bar{F}(x)} = 1 + \frac{\int_0^x \bar{F}(x-t)\, dF(t)}{\bar{F}(x)} = 1 + \int_0^y \underbrace{\frac{\bar{F}(x-t)}{\bar{F}(x)}}_{\ge 1} dF(t) + \int_y^x \underbrace{\frac{\bar{F}(x-t)}{\bar{F}(x)}}_{\ge \bar{F}(x-y)} dF(t)$$

$$\ge 1 + F(y) + \frac{\bar{F}(x-y)}{\bar{F}(x)} (F(x) - F(y))$$

Choose x large enough, s.t. $F(x) - F(y) \neq 0$ $\Rightarrow$

$$1 \le \frac{\bar F(x-y)}{\bar F(x)} \le \Big(\underbrace{\frac{\bar G(x)}{\bar F(x)} - 1 - F(y)}_{\xrightarrow[x\to\infty]{} 2}\Big) \Big/ \underbrace{(F(x) - F(y))}_{\xrightarrow[x\to\infty]{} (1-F(y))} \xrightarrow{x\to\infty} \frac{2-1-F(y)}{1-F(y)} = 1$$

$$\Rightarrow \lim_{x\to\infty} \frac{\bar F(x-y)}{\bar F(x)} = 1 \checkmark$$

ad 2): $(1) \Rightarrow \bar F(\ln y) \in \mathcal{R}$, because

$$\lim_{y\to\infty} \frac{\bar F(\ln cy)}{\bar F(\ln y)} = \lim_{y\to\infty} \frac{\bar F(\ln c + \ln y)}{\bar F(\ln y)} = 1 \checkmark$$ (Rem: y in (1) can also be < 0!)

$$\overset{\text{see above}}{\Longrightarrow} \quad y^{\varepsilon} \cdot \bar F(\ln y) \xrightarrow{y\to\infty} \infty$$

$$\overset{y=e^x}{\Longrightarrow} \quad e^{\varepsilon x} \cdot \bar F(x) \xrightarrow{x\to\infty} \infty \qquad \square$$

Remarks: 1.) (2) justifies the name „sub exponential"

2.) $\mathbb{E}[e^{\varepsilon X}] = \infty$, $\forall \varepsilon > 0$, since:

$$\mathbb{E}[e^{\varepsilon X}] \geq \mathbb{E}[e^{\varepsilon X} \mathbb{1}_{\{X \geq y\}}] \geq e^{\varepsilon y} \cdot \bar{F}(y) \longrightarrow \infty, \quad y \to \infty$$

i.e: the moment generating function does not exist!

3.) (1) is satisfied for a large class than S:

One calls these distrib. functions $\mathcal{D} \supset S$

4.) meaning of (1): for large x (and fixed y):

$\mathbb{P}(X > x)$ and $\mathbb{P}(X > x+y)$ are roughly the same!

5.) (1) implies also: ($y > 0$ fixed)

$$\frac{\mathbb{P}(X > x+y)}{\mathbb{P}(X > x)} = \frac{\mathbb{P}(X > x+y,\, X > x)}{\mathbb{P}(X > x)} = \mathbb{P}(X > x+y \mid X > x) \xrightarrow{x \to \infty} 1$$

i.e.: if X lies above a certain high barrier, then with high probability it lies also above $x+y$

6.) this is in contrast to, e.g., $X \triangleq \mathrm{Exp}(1) \notin S$

$$\mathbb{P}(X > x+y \mid X > x) = \frac{e^{-x-y}}{e^{-x}} = e^{-y} \not\longrightarrow 1$$

Generalizations to random sums

- replace $n \longrightarrow N$ r.v.

  $\mathcal{R} \longrightarrow \mathcal{S}$ $\quad \Rightarrow$

<u>Theorem</u>: Let X_i iid with distr. funct. $F \in \mathcal{S}$.

Let N an $\mathbb{N}_0$-valued r.v. indep. of the X_i,

s.t. $\mathbb{E}[z^N] < \infty$ for some $z > 1$ $\quad (*) \quad \Rightarrow$

$$\mathbb{P}(X_1 + \dots + X_N > u) \underset{u \to \infty}{\sim} \mathbb{E}[N]\,\bar{F}(u)$$

<u>Remarks</u>: a) $f(u) \overset{u \to \infty}{\sim} g(u)$ means: $\lim\limits_{u \to \infty} \frac{f(u)}{g(u)} = 1$

b) Binomial, NB, Poisson satisfy $(*)$, see Exerc.

Proof:

- we know $\overline{F}^{n*}(u) \sim n\overline{F}(u)$
- without proof: $\forall z > 1 \ \exists D < \infty$, s.t.: $\overline{F}^{n*}(u) \leq \overline{F}(u) \cdot D \cdot z^n, \ \forall u$

$$\Rightarrow \frac{\mathbb{P}(X_1 + X_2 + \dots + X_N > u)}{\overline{F}(u)} = \sum_{n=1}^{\infty} \mathbb{P}(N=n) \cdot \frac{\overline{F}^{n*}(u)}{\overline{F}(u)} \xrightarrow{u \to \infty}$$

$$\sum_{n=1}^{\infty} \mathbb{P}(N=n) \cdot n = \mathbb{E}[N].$$

Note that interchanging the $\sum$ and the limit is allowed,

because the majorant $\sum \mathbb{P}(N=n) D z^n < \infty \ \exists$ (because of (*))

$\square$

Sufficient Criterion for $F \in \mathcal{S}$ (Pitman)

Theorem:
- Let $\lambda(x) := \frac{f(x)}{\bar{F}(x)}$ be decreasing for $x \geq x_0$; $\lambda \ldots$ hazard rate
- $\lim_{x \to \infty} \lambda(x) = 0$
- $\int_0^\infty e^{x\lambda(x)} f(x)\,dx < \infty$ $\quad \Rightarrow F \in \mathcal{S}$

Proof: (for $x_0 = 0$):

Let $\Lambda(x) := \int_0^x \lambda(z)\,dz \Rightarrow \bar{F}(x) = e^{-\Lambda(x)}$, since:

$$-\lambda(z) = -\frac{f(z)}{\bar{F}(z)} = (\ln \bar{F}(z))' \quad \Big| \int_0^x dz$$

$$-\Lambda(x) = \ln \bar{F}(x) - \ln \bar{F}(0) = \ln \bar{F}(x) \checkmark$$

as above (Lemma):

$$\lambda = \frac{f}{\bar{F}} = f \cdot e^{+\Lambda}$$

$$\underline{\underline{\frac{\overline{F^{2*}}(x)}{\bar{F}(x)} - 1}} \overset{198}{=} \int_0^x \frac{\bar{F}(x-y)}{\bar{F}(x)} f(y)\,dy = \int_0^x e^{\Lambda(x) - \Lambda(x-y) - \Lambda(y)} \cdot \lambda(y)\,dy =$$

$$= \int_0^{x/2} e^{\Lambda(x) - \Lambda(x-y) - \Lambda(y)} \lambda(y)\,dy + \underbrace{\int_{\frac{x}{2}}^x e^{\Lambda(x) - \Lambda(x-y) - \Lambda(y)} \cdot \lambda(y)\,dy}$$

$$\Big| z = x - y \Big|$$

$$= \int_0^{x/2} e^{\Lambda(x)-\Lambda(x-y)-\Lambda(y)} \cdot \lambda(y)\,dy + \int_0^{x/2} e^{\Lambda(x)-\Lambda(z)-\Lambda(x-z)} \lambda(x-z)\,dz \qquad (\dagger)$$

1st integral: $\Lambda(x) - \Lambda(x-y) \underset{\lambda \text{ monotone}}{\overset{(*)}{\leq}} y \cdot \lambda(x-y) \underset{\lambda \text{ monotone},\ y < x-y}{\leq} y \cdot \lambda(y)$

$\Rightarrow$ integrand in the first integral: $\leq e^{y\lambda(y) - \Lambda(y)} \cdot \lambda(y) = e^{y\lambda(y)} f(y) \in L^1(\mathbb{R}^+)$ by assumpt.!

for fixed y: integrand $\underset{\lambda(x-y) \xrightarrow{x\to\infty} 0 \ (*)}{\longrightarrow} f(y)$

dominated convergence $\Rightarrow$ $\underline{\underline{\text{1st integral} \xrightarrow[x\to\infty]{} 1}}$

<u>2nd integral</u>:

- <u>same majorant</u>: – replace integr. variable $z \to y$!
  - $\lambda(y)$ is replaced by $\lambda(x-y)$ and: $\lambda(x-y) < \lambda(y)$ $\left(y \le \frac{x}{2}!\right)$

- <u>fixed y:</u> integrand $\le e^{y \cdot \lambda(y) - \Lambda(y)} \cdot \lambda(x-y) \xrightarrow{x \to \infty} 0$

$\overset{\text{dom. convergence}}{\Rightarrow}$ $\underline{\underline{\text{2.nd integral} \xrightarrow{x\to\infty} 0}}$

$$\overset{(+)}{=} \lim_{x\to\infty} \frac{\bar{F}^{2*}(x)}{\bar{F}(x)} - 1 = 1 \Rightarrow \lim_{x\to\infty} \frac{\bar{F}^{2*}(x)}{\bar{F}(x)} = 2 \Rightarrow F \in S \qquad \square$$

Examples for $F \in \mathcal{S}$:

- $\mathcal{R} \subset \mathcal{S}$ loggamma, Pareto $\in \mathcal{R}$
- lognormal
- heavy Weibull distr. $\bar{F}(x) = e^{-cx^{\beta}}$, $0 < \beta < 1$

Basics of Extreme value theory

Problem for ins. comp.:

- very high losses (storms, earthquakes, etc.,..)
- seldom !

tasks:

- estimation of the tails of the distr.
- calculation of risk measures (VaR, expected shortfall,..)
- premium ?

Limit distributions of Maxima

given: $X_1, \dots, X_n$ iid with distr. f. F

$$M_1 = X_1, \dots, M_n = \max(X_1, \dots, X_n)$$

Rem: results for minima follow immediately, since $\min(X_1, \dots, X_n) = -\max(-X_1, \dots, -X_n)$!

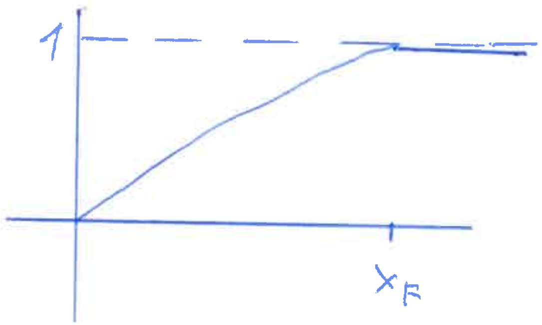
One has: $\mathbb{P}(M_n \le x) = \mathbb{P}(X_1 \le x, \dots, X_n \le x) = F^n(x)$

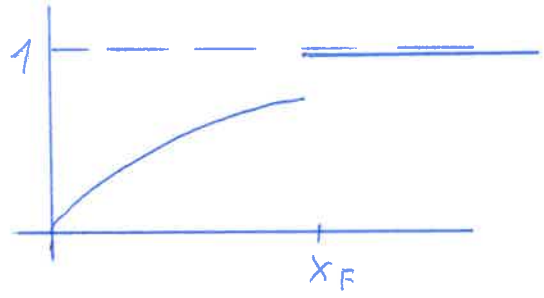
intuitively: extrema happen at the right end of the distr.!

$\Rightarrow$ "the asymptotic behavior of M_n will be determined by the behavior of F at the right end!"

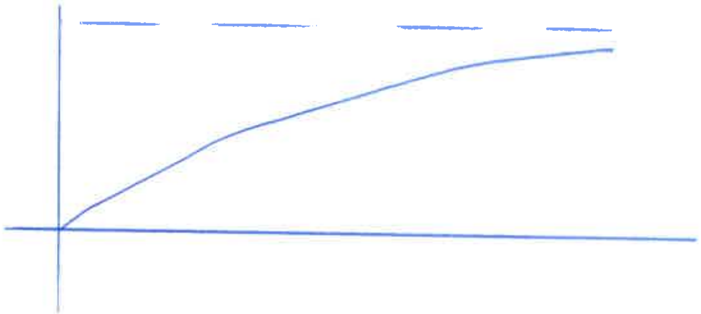
Def: $x_F := \sup\{x \in \mathbb{R} \mid F(x) < 1\}$

Ex: $x_F < \infty$





$x_F = \infty$



Obviously: $\mathbb{P}(M_n \leq x) = F^n(x) \longrightarrow 0 \quad n \to \infty \quad$ for $x < x_F$

$\Rightarrow M_n \xrightarrow{\text{in prob.}} x_F, \quad n \to \infty$

One can show $M_n \xrightarrow{a.s.} x_F, \quad n \to \infty$

not much information!

One of the central results will be:

__If__ $\exists\, c_n > 0,\ d_n \in \mathbb{R}$, s.t. $c_n^{-1}(M_n - d_n) \xrightarrow{(d)} H$ for a nondegenerated H

$\Rightarrow H$ belongs to one of 3 classes <u>(Theorem of Fisher-Tippett)</u>

Hence, we consider: $\mathbb{P}(c_n^{-1}(M_n - d_n) \le x)$

or $\mathbb{P}(M_n \le u_n(x))$

with $u_n(x) = c_n x + d_n$

Question: Which conditions on F guarantee the existence of the limit $\lim \mathbb{P}(M_n \le u_n)$ for $n \to \infty$ and appropriate u_n s?

We shall see:

- The conditions are very restrictive compared to normalized sums; there $\mathbb{E}[X^2] < \infty$ is sufficient for the central limit theorem
- important counterexample:

  The Poisson distr. has no nontrivial limit for normed maxima!

<u>Proposition</u> (Poisson Approx.): Let $\tau \in [0,\infty]$ and $u_n \in \mathbb{R}$.

Equivalent are:

(i) $n\bar{F}(u_n) \to \tau$

(ii) $\mathbb{P}(M_n \leq u_n) \to e^{-\tau}$

<u>Proof</u>: (only for $0 \leq \tau < \infty$)

<u>(i) $\to$ (ii)</u>: $\mathbb{P}(M_n \leq u_n) = F^n(u_n) = (1 - \bar{F}(u_n))^n \overset{(i)}{=} \left(1 - \frac{\tau}{n} + o\left(\frac{1}{n}\right)\right)^n \xrightarrow[n\to\infty]{} e^{-\tau}$ ✓

(ii) $\rightarrow$ (i): Claim: $\bar{F}(u_n) \rightarrow 0$ $\quad (*)$

Proof: by contradiction: assume $(*)$ is wrong $\Rightarrow \exists\, \delta > 0$ and a subsequence n_k, s.t.: $\bar{F}(u_{n_k}) \geq \delta \quad \forall k$

$$\Rightarrow \mathbb{P}(M_{n_k} \leq u_{n_k}) = (1 - \bar{F}(u_{n_k}))^{n_k} \leq (1-\delta)^{n_k} \xrightarrow{k \to \infty} 0 \quad \lightning$$

$\Rightarrow (*)$ ✓

Apply $\ln$ to (ii) $\Rightarrow$

$$n \cdot \ln F(u_n) \longrightarrow -\tau, \qquad n \to \infty$$

$$\Rightarrow -n \ln(1 - \bar{F}(u_n)) \longrightarrow \tau, \qquad n \to \infty$$

By Taylor one has: $-\ln(1-x) = x + o(x)$, $x \to 0$

$$\Rightarrow n \cdot \bar{F}(u_n) \longrightarrow \tau, \quad n \to \infty \quad \Rightarrow (i) \checkmark \quad \Rightarrow \square$$

Remarks: 1.) Let $0 < \tau < \infty$, $B_n := \sum_{i=1}^{n} \mathbb{1}_{\{X_i > u_n\}} \Rightarrow$

B_n is Binomial distr.: $B(n, \bar{F}(u_n))$!

$$\mathbb{E}[B_n] = n \cdot \bar{F}(u_n) \longrightarrow \tau \quad \approx (i) \quad : \mathbb{E}[P_o(\tau)]$$

$$\mathbb{P}(B_n = 0) = \mathbb{P}(M_n \le u_n) \longrightarrow e^{-\tau} \quad \approx (ii) \quad : \mathbb{P}(P_o(\tau) = 0)$$

Hence: „Poisson Approximation"

2.) assume, $\exists\ u_n^{(\tau)}$ for one $\tau > 0$, s.t. (i) holds true.

$\Rightarrow$ one finds $\forall \tau'$ such numbers $u_n^{(\tau')}$

Proof: (for the transition $1 \to \tau$) (instead of $\tau \to \tau'$)

Let $u_n^{(1)}$ such a sequence for $\tau = 1$.

Take $u_n^{(\tau)} := u^{(1)}_{[n/\tau]}$:

$$n \cdot \bar{F}(u_n^{(\tau)}) = \underbrace{\frac{n}{[\frac{n}{\tau}]}}_{\to\ \tau \text{ (exercise)}} \underbrace{[\tfrac{n}{\tau}] \cdot \bar{F}(u^{(1)}_{[n/\tau]})}_{\to\ 1 \text{ by assump.}} \longrightarrow \tau \qquad \square$$

One has (see above): $\mathbb{P}(M_n \leq x) \longrightarrow \begin{cases} 0 & x < x_F \\ 1 & x \geq x_F \end{cases}$.

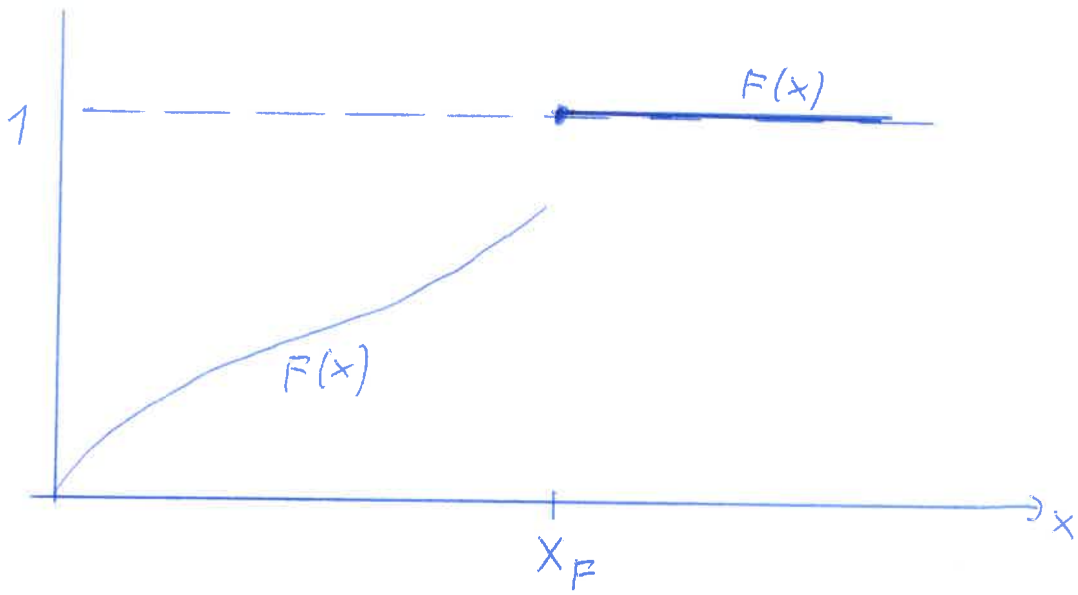
Generalization:

Lemma: assume: $x_F < \infty$ and $\bar{F}(x_F-) = F(x_F) - F(x_F-) > 0$

$\Rightarrow \forall$ sequences u_n with $\mathbb{P}(M_n \leq u_n) \longrightarrow \rho$

$\Downarrow$

$\rho = 0$ or $\rho = 1$



<u>Proof.</u> $0 \le g \le 1 \Rightarrow g = e^{-\tau}$ with $\tau \in [0,\infty]$

Proposition above $\Rightarrow$ $n \cdot \bar{F}(u_n) \xrightarrow{n\to\infty} \tau$

<u>Case 1:</u> $u_n < x_F$ for infinitely many n $\Rightarrow$

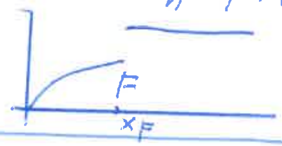
for those n: $\bar{F}(u_n) \ge \bar{F}(x_F-) > 0$

$\Rightarrow n \cdot \bar{F}(u_n) \longrightarrow \infty$, hence $\tau = \infty \Rightarrow g = 0$

<u>Case 2:</u> $u_n \ge x_F$ $\underbrace{\text{for } n \text{ large enough}}$

$\Rightarrow n \cdot \bar{F}(u_n) = 0$ $\Rightarrow \tau = 0 \Rightarrow g = 1$ □

„no matter which normalization one chooses, there does not exist a nontrivial limit distribution for the scaled M_n, if the distribution F has a jump at the finite right end.



- a similar result holds for $x_F = \infty$:

Theorem (without pr.) Let F be s.t. $x_F \leq \infty$ and $\tau \in (0,\infty)$:

$$\exists \text{ a sequence } u_n \text{ with } n\bar{F}(u_n) \longrightarrow \tau \iff$$

$$\lim_{x \to x_F} \frac{\bar{F}(x)}{\bar{F}(x-)} = 1 \quad \text{and} \quad F(x_F-) = 1$$

Special Case: (state space: $\mathbb{N}$): conditions above

$$\Big\downarrow$$

$$\lim_{n\to\infty} \frac{\bar{F}(n)}{\bar{F}(n-1)} = 1$$

<u>Example (Poisson):</u> $\mathbb{P}(X=k) = e^{-\lambda} \dfrac{\lambda^k}{k!}$

$$\frac{\bar F(k)}{\bar F(k-1)} = \frac{1-F(k)}{\bar F(k-1)} = \frac{\bar F(k-1) - \bar F(k-1) + 1 - F(k)}{\bar F(k-1)} = 1 - \frac{F(k) - F(k-1)}{\bar F(k-1)}$$

$$= 1 - \frac{\lambda^k}{k!}\cancel{e^{-\lambda}} \Big/ \sum_{r=k}^{\infty} \frac{\lambda^r}{r!}\cancel{e^{-\lambda}} =$$

$$1 - \frac{1}{1 + \underbrace{\sum_{r=k+1}^{\infty} \frac{k!}{r!}\lambda^{r-k}}} =: (*)$$

Estimation of the last sum: $= \displaystyle\sum_{s=1}^{\infty} \frac{\lambda^s}{(k+1)\cdots(k+s)} \leq \sum_{s=1}^{\infty} \left(\frac{\lambda}{k}\right)^s = \frac{\frac{\lambda}{k}}{1-\frac{\lambda}{k}} \xrightarrow{k\to\infty} 0$

$\boxed{(k > \lambda\ !)}$

$$\Rightarrow (*) \xrightarrow{k\to\infty} 0$$

$\Rightarrow$ <u>no nontrivial limit distribution for the M_n in the Poisson case</u>

Convergence of maxima under affine transformations

Consider: $c_n^{-1}(M_n - d_n) \xrightarrow{(d)} ?$

similar question: which distributions fulfill for $n \geq 2$:

$$\max(X_1, \dots, X_n) \overset{(d)}{=} c_n X + d_n , \qquad X_i \text{ iid } \sim X$$

$$c_n > 0,\ d_n \in \mathbb{R}$$

i.e.: which distributions are closed w.r.t. affine transformations?

Def: a nondegenerated r.v. X (with distr. funct. F) is called max stable, if one has: $\max(X_1, \dots, X_n) \overset{(d)}{=} c_n X + d_n$

$$c_n > 0,\ d_n \in \mathbb{R}$$

assume: X_n is a sequence with max-stable distr.

$$\Rightarrow c_n^{-1}(M_n - d_n) \overset{(d)}{=} X$$

$\Rightarrow$ $\boxed{\text{each max-stable distribution is a limit distribution of maxima for a sequence of iid r.v.}}$

Moreover, one has: these are already the only ones, because:

Theorem: $\boxed{\text{The class of max stable distributions is the same as the class of possible nondegenerated limit distributions of affine transformed maxima!}}$

Lemma (Convergence to types theorem) (without proof)

Let $A, B, A_1, A_2, \ldots$ r.v.

$b_n > 0$, $\beta_n > 0$; $a_n, \alpha_n \in \mathbb{R}$

assume: $b_n^{-1}(A_n - a_n) \xrightarrow{(d)} A$

Äquivalent are:

- $\beta_n^{-1}(A_n - \alpha_n) \xrightarrow{(d)} B$, B nondegen. (1)

- $$\left.\begin{array}{l} \lim_{n\to\infty} \dfrac{b_n}{\beta_n} = b \in [0,\infty) \\ \lim_{n\to\infty} \dfrac{a_n - \alpha_n}{\beta_n} = a \in \mathbb{R} \end{array}\right\} (2)$$

If (1) holds:

- $B = b \cdot A + a$ — a, b uniquely determined
- A (and B) are ~~then~~ nondegen. $\Leftrightarrow$ $b > 0$: (A, B belong to the same type)

Proof of the Th.: still to be shown: the possible limits of affin transf. maxima M_n are max-stable!

Let for appropriate const.: $\lim\limits_{n\to\infty} F^n(c_n x + d_n) = H(x)$, $x \in \mathbb{R}$

$$\lim_{n\to\infty} F^n(c_n x + d_n) = \mathbb{P}\left(c_n^{-1}(M_n - d_n) \leq x\right)$$

Assumption: limit distribution is continuous, s.t.: $\xrightarrow{(d)}$ $\Leftrightarrow$ pointwise converg. of the distr. funct.!

$$\left.\begin{aligned} &\Rightarrow \forall k \in \mathbb{N}: \lim_{n\to\infty} F^{nk}(c_n x + d_n) = \left(\lim_{n\to\infty} F^n(c_n x + d_n)\right)^k = H^k(x) \\ &\text{moreover: } \lim_{n\to\infty} F^{nk}(c_{nk} x + d_{nk}) = H(x) \end{aligned}\right\} \overset{\text{Lemma}}{\Longrightarrow}$$

$\exists\ \tilde{c}_k > 0,\ \tilde{d}_k \in \mathbb{R}$ with:

$$\lim_{n\to\infty} \frac{c_{nk}}{c_n} = \tilde{c}_k \quad \text{and} \quad \lim_{n\to\infty} \frac{d_{nk} - d_n}{c_n} = \tilde{d}_k$$

<u>and</u>: for $Y_1, \dots, Y_k$ iid with distr. function H:

$$\max(Y_1, \dots, Y_k) \overset{(d)}{=} \tilde{c}_k Y_1 + \tilde{d}_k$$

$\Rightarrow H$ is max stable $\square$

<u>Theorem (Fisher-Tippett)</u>: Let X_n iid. If $\exists$ constants $c_n > 0$, $d_n \in \mathbb{R}$ and a ~~den~~ nondegenerated distr. function H, s.t.

$$c_n^{-1}(M_n - d_n) \xrightarrow{(d)} H \qquad \Rightarrow$$

H belongs to one of the 3 types:

<u>Frechet:</u> $$\Phi_\alpha(x) = \begin{cases} 0 & x \le 0 \\ \exp(-x^{-\alpha}) & x > 0 \end{cases} \qquad \alpha > 0$$

<u>Weibull:</u> $$\Psi_\alpha(x) = \begin{cases} \exp(-(-x)^\alpha) & x \le 0 \\ 1 & \end{cases} \qquad \alpha > 0$$

<u>Gumbel:</u> $$\Lambda(x) = \exp(-\exp(-x)), \qquad x \in \mathbb{R}$$

idea of proof.: One has: $\forall t > 0$: $F^{[nt]}\left(c_{[nt]}x + d_{[nt]}\right) \longrightarrow H(x)$

but: $F^{[nt]}(c_n x + d_n) = \left(F^n(c_n x + d_n)\right)^{[nt]/n} \longrightarrow H(x)^t$

Conv. to types th.

$\Rightarrow$ $\exists\, \gamma(t) > 0,\ \delta(t) \in \mathbb{R}$ with

$$\lim_{n\to\infty} \frac{c_n}{c_{[nt]}} = \gamma(t) \quad \text{and} \quad \lim_{n\to\infty} \frac{d_n - d_{[nt]}}{c_{[nt]}} = \delta(t), \quad t > 0$$

and: $H^t(x) = H(\gamma(t)x + \delta(t)) \Longrightarrow$

$$\underline{H(\gamma(st)\cdot x + \delta(st))} = H^{st}(x) = \left[H(\gamma(t)\cdot x + \delta(t))\right]^s =$$

$$H(\gamma(s)[\gamma(t)\cdot x + \delta(t)] + \delta(s)) =$$

$$\underline{H(\gamma(s)\gamma(t)\cdot x + \gamma(s)\cdot\delta(t) + \delta(s))}$$

$\Longrightarrow$

$$\left.\begin{aligned} \gamma(st) &= \gamma(s)\gamma(t) \\ \delta(st) &= \gamma(s)\delta(t) + \delta(s) \end{aligned}\right\} \quad \forall s,t>0$$

one can show : these functional equations lead to the types $\Phi_\alpha, \Psi_\alpha, \Lambda$ $\square$

Remarks: 1.) The limit H is only unique up to affine transformations :

If the limit $\exists$ as $H(cx+d)$, i.e:

$$\lim_{n\to\infty} \mathbb{P}(c_n^{-1}(M_n - d_n) \le x) = H(cx+d),$$

then $H(x)$ is a possible limit as well.

$$\left(\text{choose: } \tilde{c}_n := \frac{c_n}{c} \qquad \tilde{d}_n = d_n - \frac{d\,c_n}{c} \quad , \text{ Exerc.} \right)$$

2.) Don't mix up: the present Weibull distr. Ψ_α with the ordinary one!
$\left(F_\alpha = 1 - e^{-x^\alpha}, \; x \geq 0 \right)$

3.) $\exists$ the following relations between the extreme value distributions:

X has distr. f. $\Phi_\alpha \Leftrightarrow \ln X^\alpha$ has distr. f. $\Lambda \Leftrightarrow -X^{-1}$ has distr. funct. Ψ_α

Def: The distributions $\Phi_\alpha, \Psi_\alpha, \Lambda$ are called <u>extreme value distributions</u>.

Theorem above $\Rightarrow$ <u>extreme value distr.</u> are <u>max stable</u>

One has: Frechet: $M_n \overset{(d)}{=} n^{1/\alpha} X$

Weibull: $M_n \overset{(d)}{=} n^{-1/\alpha} X$

Gumbel: $M_n \overset{(d)}{=} x + \ln n$

Example: (Exponential distr.)

$X_i \overset{d}{=} \mathrm{Exp}(1)$ iid $\Rightarrow$

$$\mathbb{P}(M_n - \ln n \le x) = (F(x+\ln n))^n = \left(1 - e^{-(x+\ln n)}\right)^n =$$

$$\left(1 - \tfrac{1}{n} e^{-x}\right)^n \xrightarrow{n\to\infty} \exp(-e^{-x}) = \Lambda(x)$$

$\Rightarrow$ "Exp (1) is in the maximum domain of attraction of the Gumbel distribution"

Maximum domain of attraction MDA and scaling constants

Question: given an extreme value distr. H, which conditions on F imply that $c_n^{-1}(M_n - d_n) \xrightarrow{(d)} H$

Similar questions:
- how to choose the scaling constants ?
- is it possible that different scaling constants lead to different limit distributions ?

$\downarrow$

Convergence to types: only up to affine transf. !

Def: X (resp. F) belongs to the maximum domain of attraction of H, if $\exists c_n > 0, d_n \in \mathbb{R}$, s.t. $c_n^{-1}(M_n - d_n) \xrightarrow{(d)} H$

one writes: $X\,(F) \in MDA(H)$

Remark: H continuous $\Rightarrow$ $c_n^{-1}(M_n - d_n) \xrightarrow{(d)} H \Leftrightarrow$

$$\lim_{n\to\infty} \mathbb{P}(M_n \le c_n x + d_n) = \lim_{n\to\infty} F^n(c_n x + d_n) = H(x), \quad \forall x \in \mathbb{R}$$

Proposition (Characterization of MDA(H)): $F \in MDA(H)$ with $c_n > 0$, d_n scaling cons.

$$\Leftrightarrow \lim_{n\to\infty} n \bar{F}(c_n x + d_n) = -\ln H(x)$$

Proof: Poisson approx.: $\Rightarrow \mathbb{P}(M_n \le u_n) \longrightarrow e^{-\tau} \Leftrightarrow n\bar{F}(u_n) \longrightarrow \tau$

set: $\left.\begin{array}{l} u_n = c_n x + d_n \\ e^{-\tau} = H(x) \end{array}\right\} \Rightarrow$

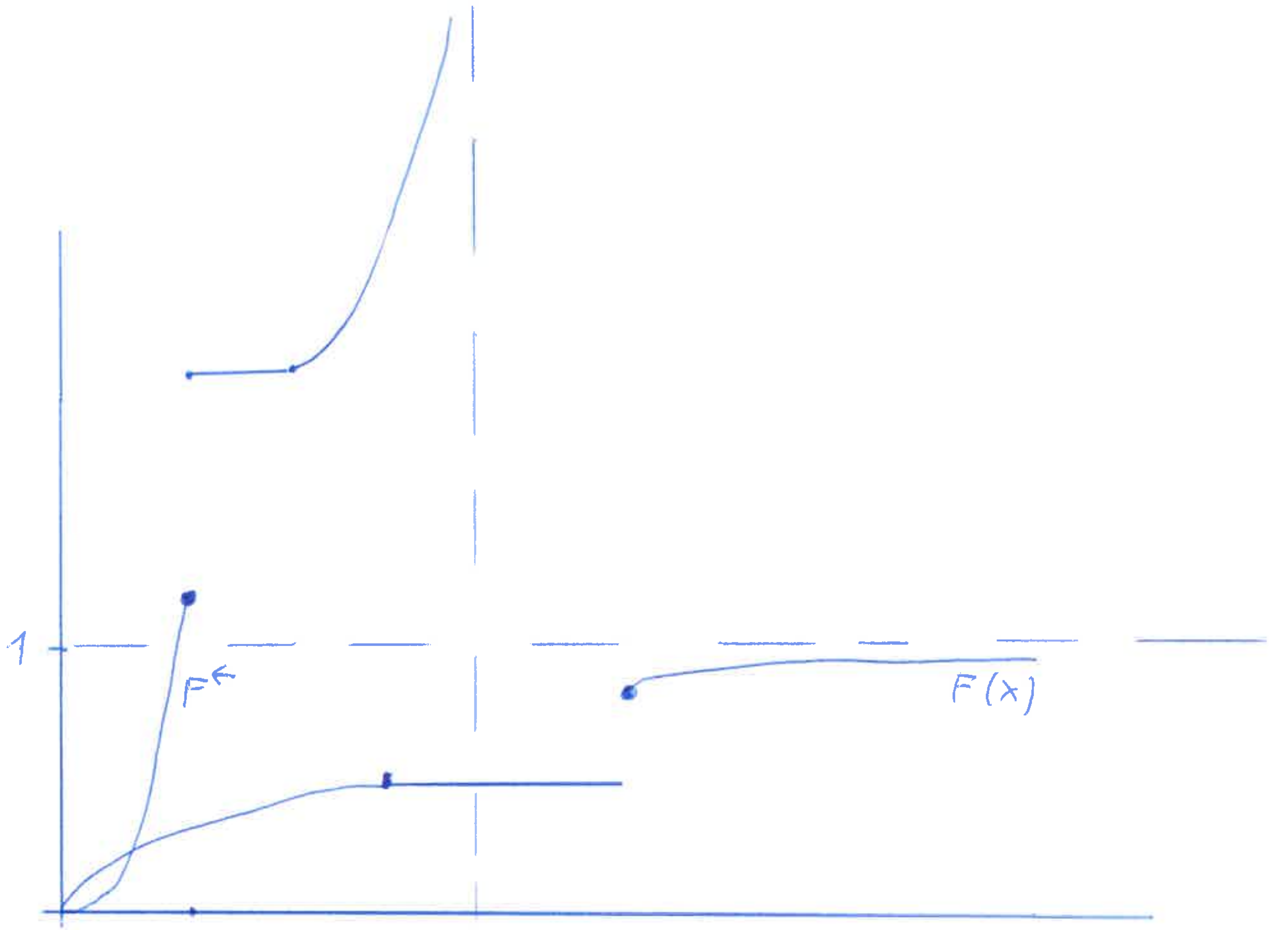
$$P(M_n \le c_n x + d_n) \longrightarrow H(x) \Leftrightarrow n\bar{F}(c_n x + d_n) \longrightarrow -\ln H(x)$$

$\square$

Def: generalized inverse of the distr. function F (quantil function)

$$F^{\leftarrow}(t) := \inf\{x \in \mathbb{R} \mid F(x) \geq t\} \qquad 0 < t < 1$$

Ex:



Maximum domain of attraction of the Frechet distribution Φ_α

Heuristics: Taylor expansion: $\bar{\Phi}_\alpha(x) = 1 - \exp(-x^{-\alpha}) \sim x^{-\alpha}$, $x \to \infty$

Question: How far away from the power are we still in $MDA(\Phi_\alpha)$?

Theorem: (without proof): $F \in MDA(\Phi_\alpha)$, $\alpha > 0 \iff \bar{F}(x) = x^{-\alpha} L(x)$, $L \in \mathcal{R}$

i.e.: $\bar{F} \in \mathcal{R}_\alpha$!

$$\text{If } F \in MDA(\Phi_\alpha) \Rightarrow c_n^{-1} M_n \xrightarrow{(d)} \Phi_\alpha \qquad (d_n \equiv 0 \,!)$$

$$c_n := F^{\leftarrow}\left(1 - \frac{1}{n}\right) = \inf\left\{x \in \mathbb{R} \,\middle|\, F(x) \geq 1 - \frac{1}{n}\right\} = \inf\left\{x \in \mathbb{R} \,\middle|\, \frac{1}{\bar{F}} \geq n\right\} = \left(\frac{1}{\bar{F}}\right)^{\leftarrow}(n)$$

Conditions on the density ??

Corollary (von Mises): Let F be absolutely continuous with density f, s.t.

$$\lim_{x\to\infty} \frac{x\, f(x)}{\bar{F}(x)} = \alpha \Rightarrow F \in MDA(\Phi_\alpha)$$

idea of proof: condition above is sufficient for $\bar{F} \in R_\alpha$ (see ex.)

Ex:

$$\left.\begin{array}{ll} \text{– Pareto:} & \bar{F}(x) = \left(\dfrac{\kappa}{\kappa + x}\right)^{\alpha} \\[2ex] \text{– Cauchy:} & f(x) = \dfrac{b}{\pi} \dfrac{1}{b^2 + (x-a)^2} \\[2ex] \text{– Burr:} & \bar{F}(x) = \left(\dfrac{\kappa}{\kappa + x^{\tau}}\right)^{\alpha} \end{array}\right\} \text{of Pareto type}$$

for these distributions one has: $\bar{F}(x) \overset{x \to \infty}{\sim} K \cdot x^{-\alpha} \Rightarrow$

$\bar{F} \in R_{\alpha} \Rightarrow F \in MDA(\Phi_{\alpha})$

scaling constants: $c_n = (Kn)^{1/\alpha}$

Maximum domain of attraction of the Weibull distr. Ψ_α

Theorem: $F \in MDA(\Psi_\alpha)$, $\alpha > 0$ $\Leftrightarrow$ $x_F < \infty$ and $\bar{F}(x_F - x^{-1}) = x^{-\alpha} L(x)$, $L \in \mathcal{R}$

If $F \in MDA(\Psi_\alpha) \Rightarrow c_n^{-1}(M_n - x_F) \xrightarrow{(d)} \Psi_\alpha$

$$c_n = x_F - F^{\leftarrow}\left(1 - \frac{1}{n}\right)$$

$$d_n = x_F$$

Remark: disadvantage in insurance appl.: $x_F < \infty$

Corollary (von Mises): If F is absolutely continuous with density f, $f > 0$ on (z, x_F) for some z, and if

$$\lim_{x \nearrow x_F} \frac{(x_F - x)\, f(x)}{\bar{F}(x)} = \alpha > 0 \;\Rightarrow\; F \in MDA(\Psi_\alpha)$$

Examples: 1.) Uniform distribution on $(0,1)$

- $x_F = 1$, $\quad F(x) = x \Rightarrow \bar{F}(x) = 1 - x \Rightarrow \bar{F}(1 - x^{-1}) = x^{-1} \in R_1$

  $\Rightarrow F \in MDA(\Psi_1)$

- using von Mises condition: $\dfrac{(1-x)\cdot 1}{1-x} = 1 \xrightarrow{x \to 1} 1$ ✓ $\Rightarrow \in MDA(\Psi_1)$
  $(f \equiv 1)$

2.) power behaviour at the right end

$\bar{F}(x) = K \cdot (x_F - x)^{\alpha}$, $\qquad$ (exercise)

Maximum domain of attraction of the Gumbel distribution $\Lambda(x)$

heuristics: Taylor: $\bar{\Lambda}(x) = 1 - e^{-e^{-x}} \sim e^{-x}, \quad x \to \infty \quad \Rightarrow$

Question: How far away from the exponential distr. are we still in $MDA(\Lambda)$??

We shall see:

- possible $x_F < \infty$

  $x_F = \infty$

- one has :
  - moderately heavy distributions (log normal, ...)
  - light distribution (normal distribution, ...)

  $\in MDA(\Lambda)$

important tool: von Mises functions!

Def: Let $x_F \le \infty$, assume $z < x_F$, s.t.:

$$\bar{F}(x) = c \cdot \exp\left(-\int_z^x \frac{1}{a(t)}\,dt\right), \quad z < x < x_F$$

where $c > 0$ is a constant, $a(t)$ pos., measurable, absolutely cont. with density a'

and $\lim_{x \to x_F} a'(x) = 0$

$\Rightarrow$ F is called von Mises function with auxiliary function a.

Ex: Exp-distr: $\bar{F}(x) = e^{-\lambda x}$, $x \ge 0$, $\lambda > 0$

Choose $z = 0$, $a(t) = \frac{1}{\lambda} \Rightarrow a' \equiv 0$ ✓ $\quad c = 1$

$$\Rightarrow \bar{F}(x) = \exp\left(-\int_0^x \lambda\,dt\right) = e^{-\lambda x} \quad ✓$$

- Weibull distribution: $\bar{F}(x) = \exp(-cx^{\tau})$, $x \geq 0$, $c, \tau > 0$, $x_F = \infty$

Choose: $a(x) = c^{-1}\tau^{-1}x^{1-\tau}$, $z = 0$

$$\Rightarrow a'(x) = c^{-1}\tau^{-1}(1-\tau)x^{-\tau} \xrightarrow[x\to\infty]{} 0 \checkmark$$

$$\bar{F}(x) = \exp\left(-\int_0^x \frac{1}{c^{-1}\tau^{-1}t^{1-\tau}}\,dt\right)$$

$$= \exp\left(-c\tau \left.\frac{t^{\tau}}{\tau}\right|_{t=0}^{x}\right) = \exp(-cx^{\tau}) \checkmark$$

- exercise: Erlang distributions, i.e. $\Gamma(n, \beta)$, $n \in \mathbb{N}$

are von Mises!

Sufficient criterion if F is C^2 at the right end of the distribution

Theorem: assume F is twice continuously differentiable in (z, x_F) for some $z < x_F$, where $F' = f > 0$ and $F''(x) < 0$ hold. $\Rightarrow$

$$F \text{ is von Mises} \iff \lim_{x \to x_F} \frac{\bar{F}(x)\, F''(x)}{f^2} = -1 \qquad (*)$$

Proof: Let $z < x < x_F$

Let $a(x) := \dfrac{\bar{F}(x)}{f(x)} > 0 \quad \Rightarrow$

$$-\frac{1}{a(t)} = -\frac{f(t)}{\bar{F}(t)} \quad \xRightarrow{\int_z^x dt}$$

$$\int_z^x -\frac{1}{a(t)}\,dt = \ln \bar{F}(t)\Big|_{t=z}^{t=x} = \ln \bar{F}(x) - \ln \bar{F}(z) \qquad \Big|\ \exp$$

$$\Rightarrow \bar{F}(x) = \bar{F}(z)\cdot \exp\left(-\int_z^x \frac{1}{a(t)}\,dt\right)$$

Moreover, $a'(x) = \dfrac{-f^2 - \bar{F}\cdot f'}{f^2} = -1 - \dfrac{\bar{F}(x)\cdot F''(x)}{f^2(x)} \Rightarrow$

(*) is equivalent to $\lim\limits_{x \to x_F} a'(x) = 0$ $\square$

Representation theorem for functions of <u>regular</u> and <u>rapid variation</u>

<u>Def:</u> a function $\bar{F}$ is of <u>rapid variation</u>, i.e. $\bar{F} \in R_\infty$, if

$$\lim_{x \to \infty} \frac{\bar{F}(xt)}{\bar{F}(x)} = \begin{cases} 0 & t > 1 \\ \infty & 0 < t < 1 \end{cases}$$

(compare: of regular variation, if limit above $= t^{-\alpha}$ („$\alpha \to \infty$" for R_∞)

Representationtheorem: A function $\bar{F}$ is of regular, (resp. rapid variation),

if we have:

$$\bar{F}(x) = c(x) \cdot \exp\left(\int_z^x \frac{\delta(u)}{u} \, du \right)$$ with c, δ measurable,

$x \geq z$ for some $z > 0$.

and $c(x) \xrightarrow{x \to \infty} c_0 \in (0, \infty)$, $\delta(x) \xrightarrow{x \to \infty} -\alpha$, (resp. $\delta(x) \xrightarrow{x \to \infty} -\infty$).

Proposition (Some properties of von Mises functions):

1.) the auxiliary function $a(x)$ can be chosen as $\frac{\bar{F}}{f}$.

2.) $x_F = \infty$: $\lim\limits_{x \to \infty} \frac{x \cdot f(x)}{\bar{F}(x)} = \infty$

3.) $x_F = \infty$: $\bar{F} \in R_\infty$

Proof: ad 1) see above

ad 2.) $\lim\limits_{x \to \infty} \frac{x \cdot f(x)}{\bar{F}(x)} = \lim\limits_{x \to \infty} \frac{x}{a(x)} = ??$

Consider: $\frac{a(x)}{x} = \frac{1}{x} \int_z^x a'(t) + \frac{a(z)}{x}$

One has: $|a'(x)| \leq \varepsilon \quad \forall x \geq N(\varepsilon) \Rightarrow$

$$a(x) < a(N) + \varepsilon(x-N), \qquad \forall x \geq N(\varepsilon)$$

$$\Rightarrow \frac{a(x)}{x} < \frac{a(N)}{x} + \varepsilon\left(1 - \frac{N}{x}\right) \qquad \Rightarrow \lim_{x\to\infty} \frac{a(x)}{x} = 0$$

$a(x)>0$

$$\Rightarrow \lim_{x\to\infty} \frac{x}{a(x)} = +\infty \quad \checkmark$$

3.) representation theorem: $\displaystyle\int_z^x \frac{\delta(u)}{u}\,du = -\int_z^x \frac{1}{a(u)}\,du \Rightarrow$

$$\delta(u) = -\frac{u}{a(u)} \xrightarrow[u\to\infty]{2.)} -\infty \checkmark \quad \square$$

Proposition (von Mises and $MDA(\Lambda)$):

Let F be a von Mises function $\Rightarrow F \in MDA(\Lambda)$.

Moreover, the scaling constants can be chosen as: $d_n = F^{\leftarrow}\left(1 - \frac{1}{n}\right)$

$c_n = a(d_n)$

Proof: Def $\longrightarrow$ for $t \in \mathbb{R}$, x close to x_F:

$$\frac{\bar{F}(x + t\,a(x))}{\bar{F}(x)} = \exp\left(- \int_x^{x + t \cdot a(x)} \frac{du}{a(u)}\right) = \left| \begin{array}{l} v = \frac{u - x}{a(x)} \\ \Rightarrow du = a(x)\,dv \end{array} \right|$$

$$= \exp\left(- \int_0^t \frac{a(x)\,dv}{a(v\,a(x) + x)}\right) \qquad (*)$$

251

We show now: integrand converges loc. unif. to 1:

Let $\varepsilon > 0$: One has $|a'(s)| < \varepsilon$ for s large enough.

Moreover,

$$\left|a(x+v\,a(x)) - a(x)\right| = \left|\int_x^{x+v\,a(x)} a'(s)\,ds\right| \underset{\substack{\uparrow \\ v \in (0,t)}}{\leq} |t| \cdot a(x) \cdot \varepsilon,$$

since a increases slower than linear, which means that the upper and lower boundary of the $\int$ are large. $\Rightarrow$

for x large enough:

$$\left|\frac{a(x+v\cdot a(x))}{a(x)} - 1\right| \leq \varepsilon |t| \Rightarrow \lim_{x \to x_F} \frac{a(x)}{a(x+v\,a(x))} = 1$$

unif. on bounded v-intervals

$$\Rightarrow \lim_{x \to x_F} \frac{\bar{F}(x + t\,a(x))}{\bar{F}(x)} = e^{-t} \quad \text{unif. on bounded } t\text{-intervals.}$$

Choose now: $x = \left(\frac{1}{\bar{F}}\right)^{\leftarrow}(n) =: d_n$

$$\Rightarrow \lim_{n \to \infty} \bar{F}(d_n + t \cdot a(d_n)) \cdot n = e^{-t}$$

Let $c_n := a(d_n) \Rightarrow \lim_{n \to \infty} n \cdot \bar{F}(d_n + t \cdot c_n) = e^{-t} = -\ln \Lambda(t)$

$$\overset{\text{Prop. above}}{\Longrightarrow} \quad F \in MDA(\Lambda) \qquad \square$$

<u>Theorem</u> (Char. of $MDA(\Lambda)$) : Let $x_F \leq \infty$

$$F \in MDA(\Lambda) \Leftrightarrow \exists\, z < x_F, \text{ s.t.}$$

$$\bar{F}(x) = c(x) \cdot \exp\left(- \int_z^x \frac{g(t)}{a(t)}\, dt \right)$$

with c, g meas.; $c(x) \xrightarrow{x \to x_F} c > 0$

$g(t) \xrightarrow[t \to x_F]{} 1$

a ... absolutely continuous with $\lim\limits_{x \to x_F} a'(x) = 0$

<u>Remarks:</u> 1.) Difference to von Mises functions: $g(t) \longleftrightarrow 1$, $c(x) \longleftrightarrow c$

2.) a growes slower than linearly!

For a ... linear: $\bar{F}$ has power behaviour!

Examples for $F \in MDA(\Lambda)$:

- Exp. distribution : $a(x) = \frac{1}{\lambda}$
- ~~Nor Ga~~ Normal distr. : $a(x) = \frac{\bar{\Phi}}{\varphi} \overset{x\to\infty}{\sim} x^{-1}$
- log normal : $a(x) \overset{x\to\infty}{\sim} \frac{\sigma^2 x}{\ln x - \mu}$
- Weibull : $a(x) = c^{-1}\tau^{-1}x^{1-\tau}$
- Benktander I : $\bar{F}(x) = \left(1 + \frac{2\beta}{\alpha}\ln x\right)\exp\left(-[\beta(\ln x)^2 + (\alpha+1)\ln x]\right)$

  $x \geq 1,\ \alpha, \beta > 0$: $a(x) = \frac{x}{\alpha + 2\beta \ln x}$
- Benktander II : $\bar{F}(x) = e^{\alpha/\beta}\, x^{-(1-\beta)} \exp\left(-\frac{\alpha}{\beta}x^{\beta}\right)$

  $x \geq 1,\ \alpha > 0,\ 0 < \beta < 1$

  $a(x) = \frac{x^{1-\beta}}{\alpha}$

Generalized extreme value and generalized Pareto distribution

Representation of Frechet, Weibull and Gumbel in a one parameter family:

<u>Def:</u> $$H_\xi(x) = \begin{cases} \exp\left(-(1+\xi x)^{-\frac{1}{\xi}}\right) & \xi \neq 0\,, \quad 1+\xi x > 0 \\ \exp(-\exp(-x)) & \xi = 0 \end{cases}$$

is called <u>generalized extreme value distribution</u>

$\xi = \alpha^{-1} > 0$ ... Frechet Φ_α

$\xi = 0$ ... Gumbel Λ

$\xi = -\alpha^{-1} < 0$ ... Weibull Ψ_α

Definition area: $\begin{cases} x > -\xi^{-1} & \xi > 0 \\ x \in \mathbb{R} & \xi = 0 \\ x < -\xi^{-1} & \xi < 0 \end{cases}$

A <u>shift parameter μ</u> and a <u>scaling parameter σ</u> provide a

<u>3 parameter family:</u>

$$H_{\xi,\mu,\sigma}(x) := H_\xi\left(\frac{x-\mu}{\sigma}\right)$$

Theorem: (Char. of $MDA(H_\xi)$): Equivalent are:

(i) $F \in MDA(H_\xi)$

(ii) $\exists$ pos., mb., function $a(u)$, s.t. for $1+\xi x > 0$:

$$\lim_{u \to x_F} \frac{\bar{F}(u + x\,a(u))}{\bar{F}(u)} = \begin{cases} (1+\xi x)^{-1/\xi} & \xi \neq 0 \\ e^{-x} & \xi = 0 \end{cases}$$

Rem: (ii) has the following probabilistic interpretation:

Let X be a r.v. with $F \in MDA(H_\xi)$

$\Rightarrow$ (ii) means:

$$\mathbb{P}\left(\frac{X-u}{a(u)} > x \,\middle|\, X > u\right) \xrightarrow{u \to x_F} \begin{cases} (1+\xi x)^{-1/\xi} & \xi \neq 0 \\ e^{-x} & \xi = 0 \end{cases}$$

$$\| $$

$$\mathbb{P}(X > u + a(u)\cdot x \,|\, X > u)$$

i.e.: right hand side is an approximation for the scaled excess over a high barrier (scaling factor: $a(u)$)

Def: Let X be a r.v. with distr. function F.

For fixed $u < x_F$:

$F_u(x) := \mathbb{P}(X-u \leq x \mid X > u)$, $x \geq 0$

is called the excess distribution of X with barrier u.

$e(u) := \mathbb{E}[X-u \mid X > u]$ is called mean excess function.

Calculation of the mean excess function:

$$\underline{\underline{e(u)}} = \frac{\int_u^{x_F} (x-u)\, dF(x)}{\bar{F}(u)} = \frac{\int_u^{x_F} (u-x)\, d\bar{F}(x)}{\bar{F}(u)} \overset{\text{integration by parts}}{=}$$

$$= \frac{(u-x)\bar{F}(x)\Big|_{x=u}^{x=x_F} + \int_u^{x_F} \bar{F}(x)\,dx}{\bar{F}(u)} = \underline{\underline{\frac{\int_u^{x_F} \bar{F}(x)\,dx}{\bar{F}(u)}}} \qquad 0 < u < x_F$$

Remark: For $u=0$ one has $(X \geq 0)$: $\mathbb{E}[X] = \int_0^{\infty} \bar{F}(x)\,dx$!

One has (see Exercises):

$$\bar{F}(x) = \frac{e(0)}{e(x)} \cdot \exp\left(- \int_0^x \bar{F}(z)\, dz \right), \qquad x > 0$$

hence: $F \longleftrightarrow e$!

Karamata Theorem (without proof):

Let $\alpha > 1$: $\int_x^\infty t^{-\alpha} L(t)\, dt \overset{x\to\infty}{\sim} (\alpha-1)^{-1} x^{-(\alpha-1)} L(x)$, $L \in \mathcal{L}$

- i.e.: one gets the asymptotic behaviour of the integral $\int_x^\infty$ over a function $\in \mathcal{R}$ in the following way: take the $L(x)$ out of the integral and integrate the remaining power as usual.

- Application: e.g.: $f \longrightarrow \bar{F}$

or:

- $\underline{\bar{F} \in R_{\alpha}\,,\ \alpha > 1}$ $\Rightarrow$ $e(u) \underset{u \to \infty}{\sim} \dfrac{u}{\alpha - 1}$

<u>Proof.</u> $e(u) = \dfrac{\int_u^\infty \bar{F}(x)\,dx}{\bar{F}(u)} = \dfrac{\int_u^\infty L(x)\, x^{-\alpha}\, dx}{\bar{F}(u)} \overset{\text{Karamata}}{\sim} \dfrac{\dfrac{\cancel{L(u)} \cdot u^{-(\alpha-1)}}{(\alpha-1)}}{\cancel{L(u)}\, u^{-\alpha}} = \dfrac{u}{\alpha - 1}$ $\square$

<u>Def:</u> The family of distr. functions G_ξ, defined by

$$G_\xi(x) := \begin{cases} 1 - (1+\xi x)^{-1/\xi}, & \xi \neq 0 \\ 1 - e^{-x} & \xi = 0 \end{cases}$$

is called <u>generalized Pareto distribution</u> (GPD)

<u>Definition area:</u>

$$\left.\begin{array}{ll} x \geq 0, & \text{for } \xi \geq 0 \\ 0 \leq x \leq -\frac{1}{\xi}, & \text{for } \xi < 0 \end{array}\right\}$$

Introducing <u>shift and scaling parameters</u>, ν, resp β $\Rightarrow$

<u>3 parameter family</u>:

$$G_{\xi,\nu,\beta}(x) := G_\xi\left(\frac{x-\nu}{\beta}\right)$$

<u>important special case</u>: $G_{\xi,0,\beta}(x) =: G_{\xi,\beta}(x) = 1 - \left(1 + \frac{\xi x}{\beta}\right)^{-\frac{1}{\xi}}$

with def. area: $D(\xi,\beta) = \begin{cases} [0,\infty), & \xi \geq 0 \\ [0, -\frac{\beta}{\xi}), & \xi < 0. \end{cases}$

Summary

- generalized extrem value distribution H_ξ describes possible limits of normalized maxima.

- generalized Pareto distribution (GPD) $G_{\xi,\beta}$ describes limit distributions of scaled excesses over high barriers.

Theorem: (properties of GPD)

a) $X \sim GPD(\xi, \beta) \Rightarrow \mathbb{E}[X] < \infty \iff \xi < 1$

one has:

- $\mathbb{E}\left[\left(1 + \frac{\xi}{\beta} X\right)^{-r}\right] = \dfrac{1}{1 + \xi \cdot r}, \quad r > -\frac{1}{\xi}$

- $\mathbb{E}\left[\left(\ln\left(1 + \frac{\xi}{\beta} X\right)\right)^{k}\right] = \xi^{k} \cdot k!, \quad k \in \mathbb{N}$

- $\mathbb{E}[X^r] = \dfrac{\beta^r}{\xi^{r+1}} \dfrac{\Gamma(\xi^{-1} - r)}{\Gamma(1 + \xi^{-1})} r!, \quad \xi < \frac{1}{r}, \ r \in \mathbb{N}$

b) $\xi \in \mathbb{R}$, $F \in MDA(H_\xi) \iff$

$$\lim_{u \to x_F} \sup_{0 < x < x_F - u} \left| F_u(x) - G_{\xi, \beta(u)}(x) \right| = 0$$

for a positive function $\beta(u)$

c) $X \triangleq GPD(\xi, \beta)$, $\xi < 1 \iff$ for $u < x_F$

$$e(u) = \mathbb{E}[X - u \mid X > u] = \frac{\beta + \xi u}{1 - \xi}$$

idea of proof: a) calculation

b) compare Th. above: there the excess is ~~scal~~ scaled, hence

we have in our case

$G_{\xi, \beta(u)}$ as limit!

c) calculation using $e(u) = \frac{1}{\bar{F}(u)} \int_u^{x_F} \bar{F}(x)\,dx$

<u>Remark:</u> The theorem motivates the following heuristic procedure („<u>peak over threshold</u>" (POT) <u>method</u>)

- in areas where the GPD is a good approximation for the cond. distribution: $e(u)$ should be approximately linear $\Rightarrow$
- determine where <u>$e(u)$ is approx. linear</u> by plotting the <u>empirical mean excess function</u>

Let this be the case for $u \geq u_0$

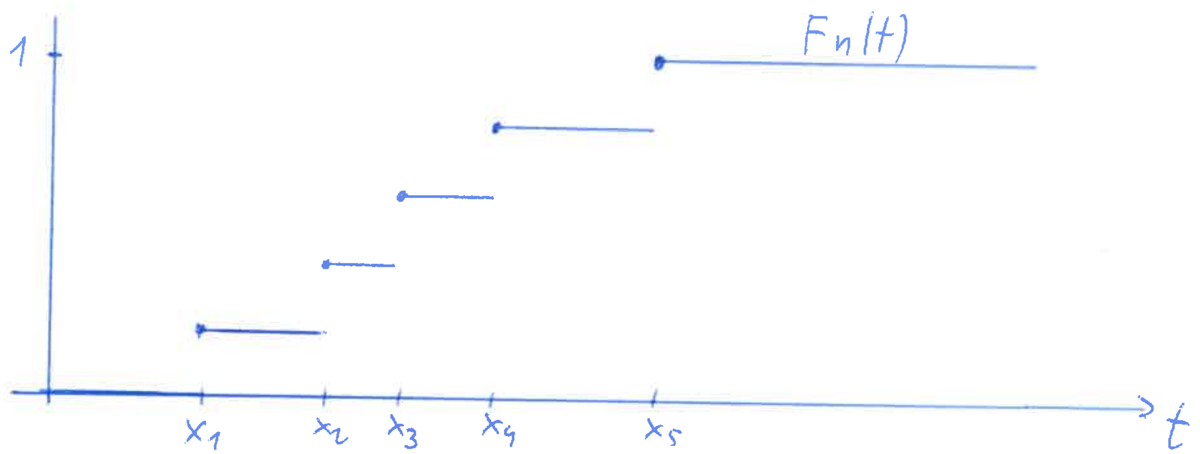
- for $0 \leq u \leq u_0$: use standard methods (enough data)
- for $u > u_0$: use the GPD approximation!

Example: (Calculation of the p-quantil, i.e the VaR at level p) e.g: $p = 0.99$

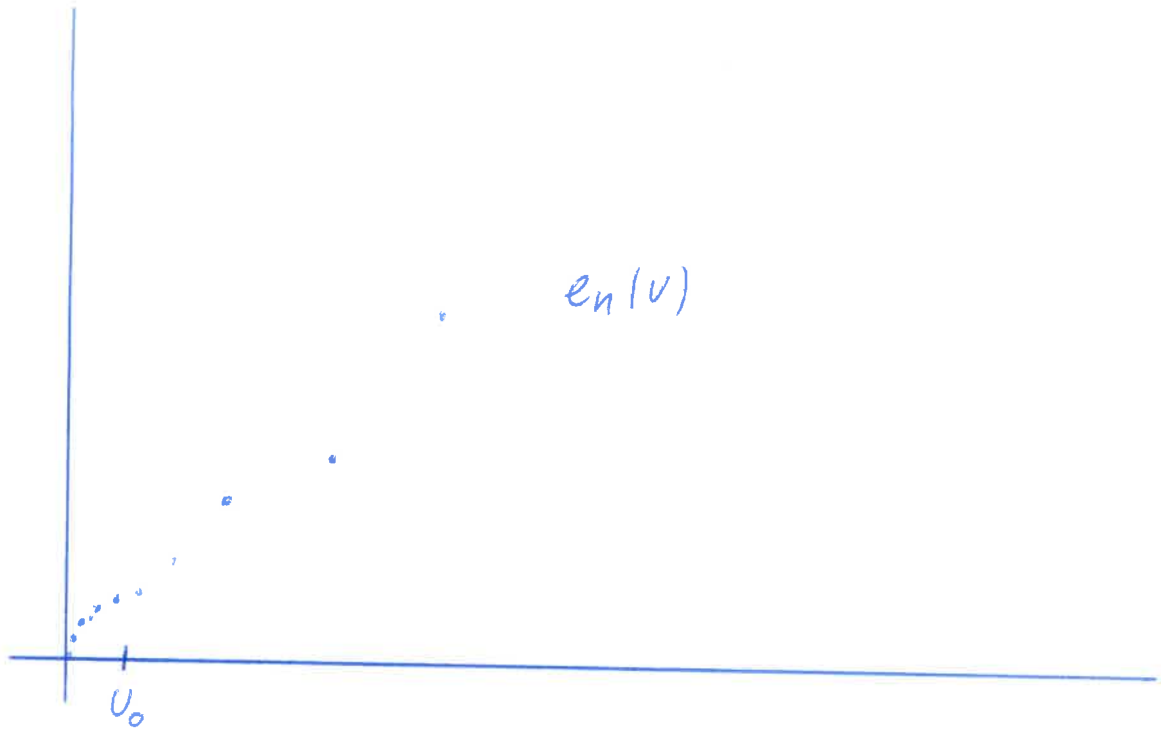
Def: empirical mean excess function

$$e_n(u) := \frac{1}{\bar{F}_n(u)} \int_u^\infty \bar{F}_n(y)\, dy$$

$$F_n(t) := \frac{1}{n} \sum_i \mathbb{1}_{\{x_i \le t\}}$$ ... emp. distribution function



assume e_n looks like:



Write now, for convenience, v instead of v_0!

Procedure: 1.) $\hat{\bar{F}}(u) = \bar{F}_n(u) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}\{X_i > u\} =: \frac{N_u}{n}$

2.) $\widehat{\bar{F}_u}(y) = \bar{G}_{\hat{\xi},\hat{\beta}}(y)$ (ML method for $\hat{\xi}, \hat{\beta}$)

3.) $\widehat{\bar{F}(u+y)} = \frac{N_u}{n} \cdot \left(1 + \hat{\xi}\frac{y}{\hat{\beta}}\right)^{-1/\hat{\xi}}$

Quantil ? :

Solve: $(1-p) = \frac{N_u}{n}\left(1+\hat{\xi}\frac{y}{\hat{\beta}}\right)^{-1/\hat{\xi}}$ for y ! $\Rightarrow$

$$\left[(1-p)\cdot\frac{n}{N_u}\right]^{-\hat{\xi}} = 1 + \hat{\xi}\frac{\hat{y}}{\hat{\beta}}$$

$$\Rightarrow \hat{y} = \left(\left[(1-p)\frac{n}{N_u}\right]^{-\hat{\xi}} - 1\right)\cdot\frac{\hat{\beta}}{\hat{\xi}} \Rightarrow \hat{x}_p = u + \hat{y}$$

„Real world example"

- cooperation with BANK AUSTRIA: calculation of risk measures for operational risk

- Operational risk: short definition: risk of a bank, not included in credit or market risk (e.g. legal costs, fraud, computer troubles)

- typical: heavy-tailed distributions!

  $\Rightarrow$ risk measure: VaR

Problem: given: • historical data of the BA

- extern data of the ORX bank association
- data classified with respect to „businesslines" and „types of events"

goal: VaR for the individual business lines

methods:
- extreme value theory (heavy-tailed data!)
- Bayes methods (internal data + ORX data)
- collective model

275

Combining Extreme Value Theory with Credibility Methods and Point Processes for Modelling Operational Risk

Grigory TEMNOV

Abstract

We investigate the classical problem of operational risk estimation. The usual measure for operational risk — Value-at-Risk is used, applied to predicted future yearly aggregate losses. According to it, the problem is reduced to the modelling of yearly aggregate loss for the selected Bank, given the historical database of occurred operational losses.

The loss data is classified with respect to the usual division to lines of business and to types of events. Besides the internal database, containings records of losses accumulated by the Bank, the external database, which is the collection of the data from different banks, is available. This sort of data presentation provides the first problem to be solved — how to mix internal and external data properly in order to obtain most reliable model for the loss occurrence process. We approach this problem using the full bayesian methodology.

As soon as the models for the single loss distributions are built, the problem of calculation of aggregate loss distribution characteristics appears. To obtain VaR estimation, and also some other useful characteristics of the aggregate loss distribution, such as expected shortfall, we use three different approaches, each of which appears to have its own advantages, becoming important in different variations of the problem. The first method we use is Monte-Carlo simulations, the second one bases on the application of Fourier transform, which we use by adopting Fast Fourier Transformation scheme, and the third approach is based on the use of probability generation function for aggregate loss in the form of power series and uses the recursive method to [illegible]ate coefficients in this power series representation.

[illegible]ompare the results obtained with each of three methods, applying them to different segments of data, and analyze efficiency of the three corresponding algorithms.

The Problem

GIVEN	TO BE FOUND
• Database of historical internal OpLoss data from selected bank	α–VaR for yearly aggregate losses $\mathrm{VaR}_\alpha(S) = \inf\{ s \in \mathbf{R} : \mathrm{P}(S > s) \leq 1-\alpha \}$
• External OpLoss data from "ORX" bank association	• Capital charge for OpRisk must be calculated as $C = \sum_{l=0}^{K} VaR_\alpha^l$,
• All the data classified w.r.t. *Lines of Business* (BL) and *Types of Events* (ET)	where $l = 1, \dots, K$ denotes BL's.

METHODS

- Severity distributions for BL's are very heavy tailed → EVT is used
- Homogeneous and inhomogeneous Poisson processes (and Negative Binomial distribution are used to model loss frequency)
- To evaluate aggregate loss several methods are used: *Monte-Carlo modelling, FFT approach, Recursion formulae*

Fitting severity distributions

Generalized Pareto distribution (GPD)

$$G_{\xi,\mu,\beta}(x) = 1 - \left(1 + \xi\frac{x-\mu}{\beta}\right)^{-1/\xi} \tag{1}$$

μ location parameter (threshold), β scale parameter, ξ shape parameter.

The density for GPD

$$g_{\xi,\mu,\beta}(x) = \frac{1}{\beta}\left(1 + \xi\frac{x-\mu}{\beta}\right)^{-\frac{1}{\xi}-1} \tag{2}$$

Log–likelihood function for MLE parameter estimation

$$l_g(\theta) = L_g((\xi,\beta);\mathrm{X}) = -n\ln\beta - \left(\frac{1}{\xi}+1\right)\sum_{j=1}^{n}\ln\left(1 + \frac{\xi}{\beta}(X_j-\mu)\right) ; \tag{3}$$

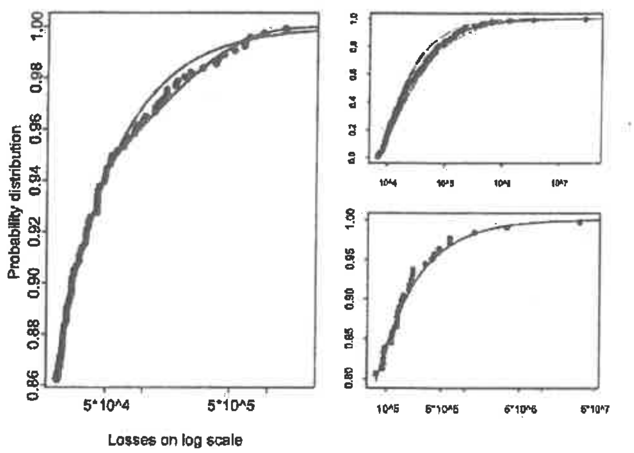
Three–parameter Weibull distribution

$$W_{\xi,\mu,\beta}(x) = 1 - e^{\left(\frac{-(x-\mu)}{\beta}\right)^{\xi}}, \tag{4}$$

$$\text{density: } w_{\xi,\mu,\beta}(x) = \xi\,\beta^{-\xi}\,x^{\xi-1}\,e^{|-(x-\mu)/\beta|^{\xi}}. \tag{5}$$

$$\text{Log–likelihood: } l_w(\theta) = \ln\xi - \xi\ln\beta + \sum_{i=1}^{n}\left((\xi-1)\ln X_i - \left[\frac{X_i-\mu}{\beta}\right]^{\xi}\right) \tag{6}$$

Figure 1. Fit with analytical distributions for selected BL's,
Top right: BL 10 — GPD fit with error bounds, bottom right — tail of BL 10 fit,
Left: BL2 — Combination of GPD (brown) and Weibull (green) — for the tail



Mixing internal and external data

Bayesian inference

$$\pi_{\theta\,|\,X}(\theta\,|\,\mathbf{x}) = \frac{f_{X\,|\,\theta}(\mathbf{x}\,|\,\theta)\pi(\theta)}{\int f_{X\,|\,\theta}(\mathbf{x}\,|\,\theta)\pi(\theta)d\theta}, \tag{7}$$

$\pi_{\theta\,|\,X}(\theta\,|\,\mathbf{x})$ — posterior parameter distribution

$\pi(\theta)$ — prior parameter distribution

For the case of **GPD** and **Poisson** joint model

- Prior for Poisson intensity λ — gamma distribution $\gamma_{a,b}(\lambda)$
- Prior for shape parameter ξ — gamma distribution $\gamma_{c,d}(\xi)$
- Prior for scale parameter β — reciprocal gamma $p(\beta) = \mathrm{P}(B=\beta)$ $(\, \mathrm{P}(1/B = \beta) = \gamma_{e,f}(\beta)\,)$

Then Posterior density for scale parameter β:

$$\tilde{p}(\beta) = \beta^{-k}(d'(\beta))^{-(a+k)}\exp\left(-\sum_{i\leq k}\log(1 + X_i/\beta)\right)p(\beta),$$

where

$a' = a+k$, $b' = b+T$, $c' = c+k$ and $d'(\beta) = d + \sum_{i\leq k}\log(1+X_i/\beta)$

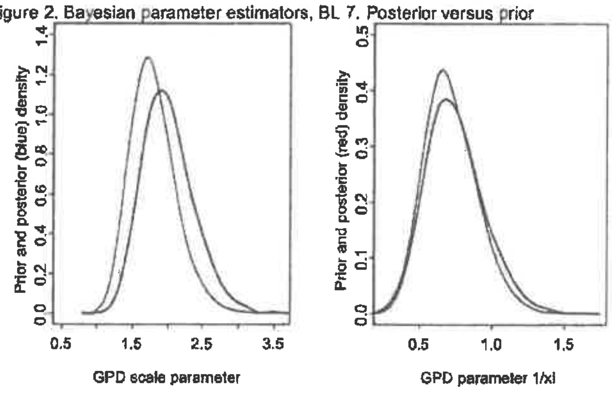
Bayesian estimators for λ, ξ and β

$$\lambda^* = \frac{a+k}{b+T} \tag{8}$$

$$\xi^* = \int \frac{c'}{d'(\beta)}\tilde{p}(\beta)d\beta \tag{9}$$

$$\beta^* = \int \beta\tilde{p}(\beta)d\beta \tag{10}$$

Figure 2. Bayesian parameter estimators, BL 7. Posterior versus prior



TECHNISCHE UNIVERSITÄT WIEN
VIENNA UNIVERSITY OF TECHNOLOGY

Financial and Actuarial Mathematics Research Group
Vienna University of Technology

for the 31st Conference on Stochastic Processes and their Applications, 17 July 2006

Claims Reserves

Examples:

- premium reserves
- reserves for large claims
- late claims reserves

Remarks:

- reserves have to be estimated
- often higher than the total premiums / year
- too high $\Rightarrow$ security $\nearrow$, but profit $\searrow$ $\Rightarrow$ taxes $\searrow$

  $\Rightarrow$ legal regulations !

- too low $\Rightarrow$ profit $\nearrow$ but security $\searrow$

<u>premium reserves:</u> e.g. premiums paid by instalments

<u>Reserves for very high claims:</u>

<u>Ex:</u>

year ↓	#	total loss in million €
	2110	32.0
	1907	28.4
	<u>2083</u>	<u>70.2</u> $\longrightarrow$ reason: one claim with 40 million € !
	2021	30.5
	2150	32.8

Late claims reserves

- occurence of the claim $\longrightarrow$ regulation: often a couple of years

  - late manifestation (e.g. dangerous substances in food, ..)
  - long regulation duration (lawsuits, expert opinions, ...)

INBR ... incurred but not reported
RBNS ... reported but not settled

We distinguish:

- <u>deterministic models</u> (basically one estimates the $\mathbb{E}$!)
- <u>stochastic models</u> (variances and distributions are also considered)

– micro models : models for the individual claims, then $\sum$

– macro models : start with the whole portfolio

IBNR methods – deterministic approach

notation: X_{ij} … non cumulative values, i.e.

payments for claims in year i regulated in year j.

or cumulative values, i.e.

payments for claims in year i, regulated until year j.

$$1 \le i, j \le t$$

Example:

non cumulative

Year ↓ \ regulation year	1	2	3	4	5	6
2017	234	105	37	21	14	2
2018	255	107	39	35	13	
2019	215	98	35	20		
2020	270	125	45			
2021	262	119				
2022	265					

$1 \le i, j \le 6 = t$

$X_{3,2} = 98$ claims from 2019

regulated 2020

cumulative

Year ↓	1	2	3	4	5	6
2017	234	339	376	397	411	413
2018	255	362	401	436	449	
2019	215	313	348	368		
2020	270	395	440			
2021	262	381				
2022	265					

$X_{3,2} = 313$

claims from 2019, regulated

2019 or 2020

known: $X_{i,j}$ with $i+j \le t+1$

claims triangle

goal: estimate the $X_{i,j}$ in the lower triangle, i.e. $i+j > t+1$

notation: total claims of the year i: $X_i = \lim\limits_{"s \to \infty"} X_{i,s}$ in the cumulative Δ

reserves for the claims in year i: $Y_i = X_i - X_{i,t+1-i}$

(paid until now: $X_{i,t+1-i}$!)

Assumptions: • everything is regulated after t years, i.e.

$$X_i = X_{i,t}$$

• no break in trends in the last years (e.g. changement in jurisdiction, ...)

• heuristic method to estimate the quality of certain deterministic methods

– remove the diagonal $X_{1,t}, X_{2,t-1}, \dots X_{t,1}$

– apply the method to the reduced Δ $\Rightarrow \hat{X}_{1,t}, \hat{X}_{2,t-1}, \dots, \hat{X}_{t-1,2}, \hat{X}_{t,1}$

– compare with the data $X_{1,t}, \dots$

Chain ladder method

- one of the oldest methods
- Assumption: in the cumulative Δ : columns are proportional, i.e.

$c_{s,s'}$ ... factor to come from column s to s'.

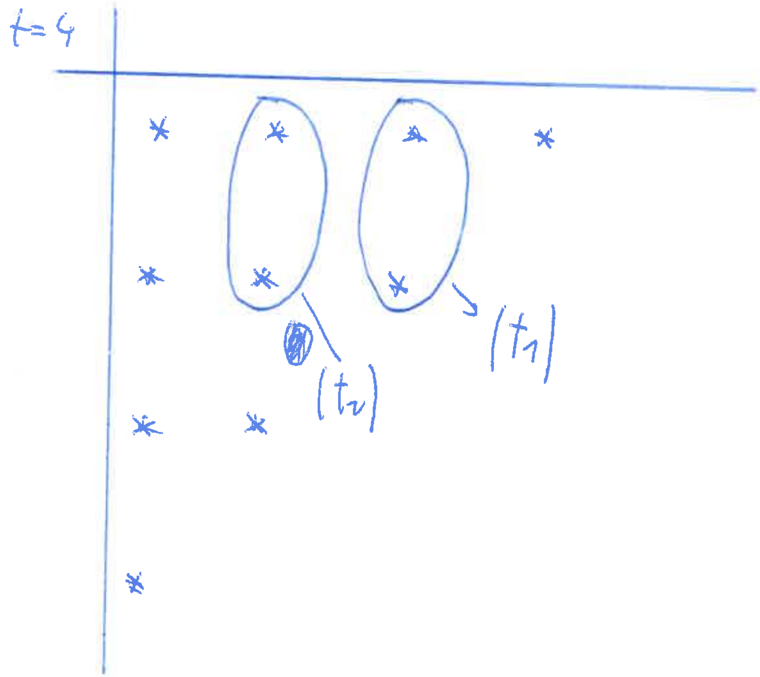
settlement factor

$$\Rightarrow X_{i,s'} \text{ "="} \; c_{s,s'} X_{i,s} \,, \qquad s' > s$$

$$\Rightarrow c_{s,s'} = \prod_{n=s}^{s'-1} c_{n,n+1} \,, \qquad c_s := \text{"} c_{s,\infty} \text{"} = \prod_{n \geq s} c_{n,n+1}$$

hence, it is sufficient to estimate the $c_{n,n+1}$!:

<u>classical chain ladder:</u> $$\hat{c}_{s,s+1} = \frac{\sum_{n=1}^{t-s} X_{n,s+1}}{\sum_{n=1}^{t-s} X_{n,s}}$$



i.e.: $c_{2,3}$ "=" $\frac{(t_1)}{(t_2)}$

$$\hat{X}_{i,s} = \hat{c}_{t+1-i,s} \cdot \underbrace{X_{i,t+1-i}}_{\text{element of the diagonal}} \qquad \text{for } s > t+1-i$$

e.g.: $\hat{X}_{3,3} = \hat{c}_{2,3} \cdot X_{3,2}$

<u>estimator for the total reserve for year i</u>

$$\hat{Y}_i = \hat{X}_{i,t} - X_{i,t+1-i} = (\underbrace{\hat{c}_{t+1-i}}_{= \prod_{n \geq t+1-i} c_{n,n+1}} - 1) X_{i,t+1-i}$$

- often: an additional security factor $(1+\alpha)$, $\alpha > 0$.

disadvantages:

- $X_{i,t+1-i} = 0 \Rightarrow \hat{Y}_i = 0$

  (if there are no claims until now $\Rightarrow$ no reserves ☹)

- sensitive w.r.t. changes in the data.

London - chain ladder

- <u>chain ladder</u>: $X_{j,s+1} = c_{s,s+1} X_{j,s}$

  geometric: straight line through the origin

- <u>London chain ladder</u>: <u>general straight line</u>: $X_{j,s+1} = b_s X_{j,s} + a_s$

  or: $X_{j,s+1} = q_s (X_{j,s} - p) + p$

  <u>rotation</u> around (p,p)

Estimator: calculate <u>regression line</u>: $\sum_{j=1}^{t-s} \left[X_{j,s+1} - (b_s X_{j,s} + a_s) \right]^2 \longrightarrow \min$

(for the first ansatz)

for the 2nd ansatz:

- choose a pivot point (p_0, p_0)

- estimate q_s by:

$$\min \sum_{j=1}^{t-s} \left(X_{j,s+1} - [q_s (X_{j,s} - p_0) + p_0]\right)^2 \quad \Rightarrow \hat{q}_s$$

- new pivot point (p_1, p_1) by:

$$\min \sum_{s=1}^{t-1} \sum_{j=1}^{t-s} \left(X_{j,s+1} - [\hat{q}_s (X_{j,s} - p) + p]\right)^2$$

- iteration!

D-chain ladder

- use cumulative values

- set $D_{i,j} = \dfrac{X_{i,j+1}}{X_{i,j}}$, $\quad i+j \leq \underline{\underline{t}}$

$\Rightarrow \Delta$:

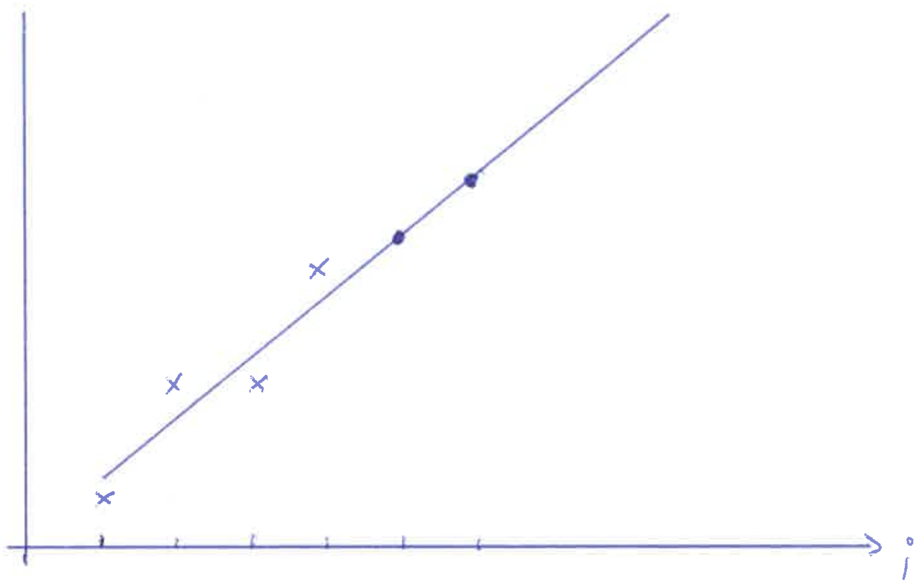
	1	2	3	. . .	t-1
1	$D_{1,1}$	$D_{1,2}$	$D_{1,3}$		$D_{1,t-1}$
2	$D_{2,1}$	$D_{2,2}$			
3					
⋮					
t-2	$D_{t-2,1}$	$D_{t-2,2}$			
t-1	$D_{t-1,1}$				

estimators for $D_{i,j}$ $\quad i+j>t$

- last column: only $D_{1,t-1}$ is known $\Rightarrow$ $\hat{D}_{i,t-1} = D_{1,t-1}$
- last but one column: $\hat{D}_{i,t-2} = \dfrac{D_{1,t-2}+D_{2,t-2}}{2}$
- other columns:

e.g.

x ... given



- estimators by regression line

$\longrightarrow$ completion of the D matrix

$\longrightarrow$ back transformation to $X_{i,j}$

multiplicative model

- Assumption: $\exists\ \theta_i > 0$, $1 \le i \le t$, s.t. one has for the non-cumulative values

$$\mathbb{E}[X_{i,j}] = \alpha_i \theta_j$$

α_i ... expected total loss in year i,

hence $\alpha_i = \mathbb{E}[X_{i,1} + \dots + X_{i,t}] = \alpha_i \theta_1 + \dots \alpha_i \theta_t$

$$\Rightarrow \sum_{j=1}^{t} \theta_j = 1$$

interpretation: θ_j ... fraction of the total loss generated in year j

Estimation of the α_i, θ_j (de Vylder)

$$\sum_{i+j \le t+1} \left(X_{i,j} - \hat{\alpha}_i \hat{\theta}_j \right)^2 \longrightarrow \min \quad (*), \quad \text{constraint: } \sum_{j=1}^{t} \hat{\theta}_j = 1$$

Setting the derivatives equal to zero $\Rightarrow \ldots \Rightarrow$

$$\hat{\alpha}_i = \frac{\sum_{j \le t+1-i} X_{i,j} \hat{\theta}_j}{\sum_{j \le t+1-i} \hat{\theta}_j^2} \qquad \hat{\theta}_j = \frac{\sum_{i \le t+1-j} X_{i,j} \hat{\alpha}_i}{\sum_{i \le t+1-j} \hat{\alpha}_i^2}$$

- nonlinear system
- iterative solution with initial value, e.g., $\hat{\theta}_j = \frac{1}{t}$
- if one has „convergence": finally normalization:

$$\left.\begin{aligned} \hat{\theta}_j &\longrightarrow \lambda \hat{\theta}_j =: \hat{\theta}_j^{new} \\ \hat{\alpha}_i &\longrightarrow \frac{1}{\lambda} \hat{\alpha}_i =: \hat{\alpha}_i^{new} \end{aligned}\right\} (+)$$

$$\text{s.t.: } \sum \hat{\theta}_j^{new} = 1$$

Reason: (*) is invariant w.r.t. the transformation (+) !

method of marginal sums

idea: in the Δ:
- summation over each column and each row
- equate to the multiplicative ansatz !

i.e.:
$$\sum_{i=1}^{t+1-j} X_{i,j} = \sum_{i=1}^{t+1-j} \hat{\alpha}_i \hat{\theta}_j \qquad j = 1, \ldots, t \qquad (1)$$

$$\sum_{j=1}^{t+1-i} X_{i,j} = \sum_{j=1}^{t+1-i} \hat{\alpha}_i \hat{\theta}_j \qquad i = 1, \ldots, t \qquad (2)$$

motivated by:
$$\sum_{i=1}^{t+1-j} \mathbb{E}[X_{i,j}] = \sum_{i=1}^{t+1-j} \alpha_i \theta_j \qquad j = 1, \ldots, t$$

$$\sum_{j=1}^{t+1-i} \mathbb{E}[X_{i,j}] = \sum_{j=1}^{t+1-i} \alpha_i \theta_j \qquad i = 1, \ldots, t$$

in the multiplicative model !

without proof.: • (1) (2), $\sum_{j=1}^{t} \hat{\theta}_j$ have exactly one solution

• the estimated total loss $\hat{\alpha}_i$ coincides with the classical chain ladder