IAA AAI

SECTION COLLOQUIUM 2019

THE MODERN ACTUARY - CHALLENGE • INFLUENCE • LEAD

ASTIN • IAAHS • IAALS • IACA • PBSS

2 - 5 April

2019

Cape Town
South Africa

CTICC

www.colloquium2019.org.za

Hosted by

ACTUARIAL SOCIETY OF SOUTH AFRICA

A Quantile Mixing Approach for the Combination of Experts' Models

Thierry COHIGNAC (*) & Nabil KAZI-TANI (**) France

(*)

(**)

Introduction

Consider that we have the opinion of two experts (two models) for a given phenomenon:

- How to build a model from the two initial models ?
- How to combine these models for risk management and / or pricing ?

Classical Actuarial Technique: Theory of Credibility (1/4)

Goal: estimation of a parameter associated with a risk X (frequency and/or cost) belonging to a class of homogeneous risks

Hypothesis

- We observe a sample of size n of the variable X,
- We denote $\bar{X}$ the empirical mean of the sample
- The sample belong to a class of risks
- The average of the risk class is denoted "μ",

- We denote $\hat{\mu}$ the mean estimate of X (individual risk)

Result (Bühlmann & all)

$$\hat{\mu} = w \times \bar{X} + (1 - w) \times \mu$$

With w, the weight (credibility) given to the sample mean

Classical Actuarial Technique: Theory of Credibility (2/4)

« W » computation (Bayesian approach)

$$W = \frac{n}{n + K}$$

With

– n, sample size (when n –> ∞, w–> 1),

– K = $\frac{EPV}{VHM}$,

– EPV : Expected Value of Process Variance,

– VHM : Variance of the Hypothetical Mean

Remarks:

- « W » is obtained by minimizing the variance of the estimator $\hat{\mu}_i$
- Total variance theorem: Var(Y) = EX(Var[Y/X] + VarX([E(Y/X]) = EPV + VHM

Classical Actuarial Technique: Theory of Credibility (3/4)

Examples of applications of credibility theory

1) Frequency computation for a particular risk belonging to a "homogeneous" group of risks (ie motor insurance)

Hypothesis (model Poisson-Gamma)

- Losses number for year j, N_j ~ Poisson law with θ parameter
- The parameter θ ~ Gamma law with (γ, β) parameters

Then

$$F^{bayes} = \alpha \times F^{ind} + (1 - \alpha) \times F^{col}$$

With

- $F^{ind} = \overline{N}$: individuel frequency of the individual risk
- $F^{col} = \frac{\gamma}{\beta}$: (group frequency equal to the gamma Gamma distribution mean)
- $\alpha = \frac{n}{n+K}$, K = β

Examples of applications of credibility theory

2) Computation of the melting frequency of a nuclear rector

Main hypothesis: Binomial-Beta model for frequency

- The accident number fallows a Binomial law of parameter p (deduced from the observed frequency of nuclear accidents « $F^{observed}$ »)
- the p parameter fallows a Beta law with (st, (1-st)) parameters, (t corresponds to the frequency calculated by the experts "Fexpert" in nuclear safety, s to the credibility granted to the a priori)

Then

$$F^{bayes} = \alpha \times F^{observed} + (1 - \alpha) \times F^{expert}$$

With

- $\alpha = \frac{n}{n+s}$,
- n: number of years reactor (knowing that there are about 500 reactors with an average seniority of 28.8 years).
- s: strength of the prior

Reference: The Economics and Uncertainties of Nuclear Power François Lévêque Ed Cambridge

From Distribution parameters to entire Distribution

Credibility theory give us a parameter, for example the mean (useful for pricing), but not the entire distribution mixture law…

Models combination

Reference Publications :

- Bates and Granger, The Combination of Forecasts, **1969**.

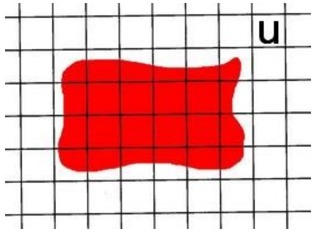
- Granger, Combining Forecasts: twenty years later, **1989**.

Granger : 2003 Nobel Price for economics

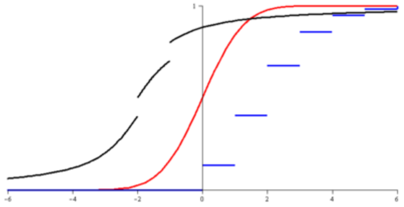
Models combination

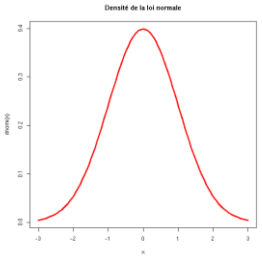
Common method of previous papers: Linear combination of

- Mesures

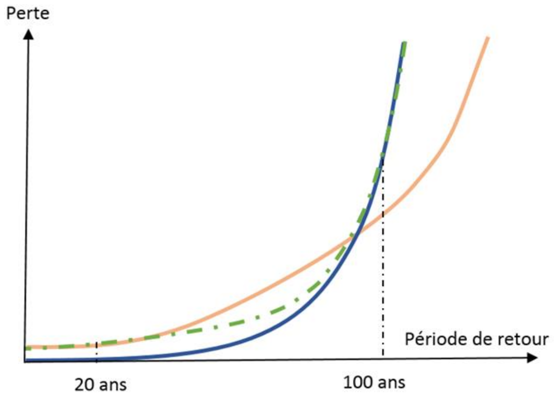


- Distribution Function or densities





- Quantiles



Quantile Combination

Beetwen these possibilities we choose « quantile linear combination »

What is needed ?

- A dependency structure (joint law)
- Weights for each quantile

Quantile Combination

Step 1 : joint law

Joint law is chosen in order to minimize disagreement between the two « models » (or experts) represented by two mesuress μ and ν

Let:

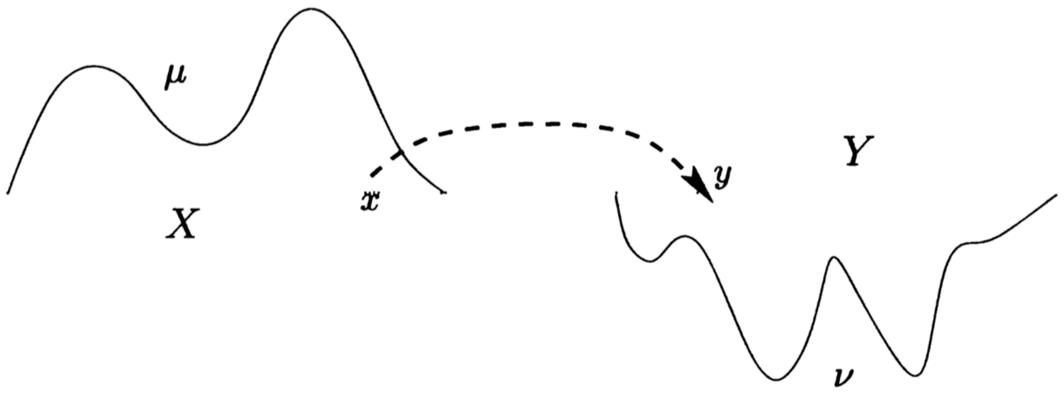
- $\pi \in \Pi(\mu, \nu)$: probability measure on R^2 with marginal given by μ and ν
- π is chosen in order to minimise the disagreement cost (models disagreement) :

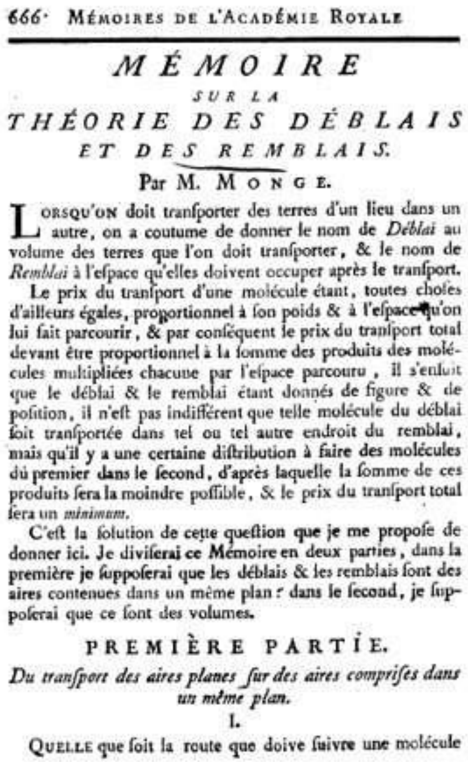
$$V_c := \inf_{\pi \in \Pi(\mu,\nu)} \int_{\mathbb{R}^2} c(x,y) d\pi(x,y).$$

This is an « optimal transport » formulation of the problem
Reference: C Villani, Book, topics in optimal transportation

The optimal transportation theory origine….

Monge (1746-1818)





666· MÉMOIRES DE L'ACADÉMIE ROYALE

MÉMOIRE
SUR LA
THÉORIE DES DÉBLAIS
ET DES REMBLAIS.
Par M. MONGE.

LORSQU'ON doit tranſporter des terres d'un lieu dans un autre, on a coutume de donner le nom de *Déblai* au volume des terres que l'on doit tranſporter, & le nom de *Remblai* à l'eſpace qu'elles doivent occuper après le tranſport.

Le prix du tranſport d'une molécule étant, toutes choſes d'ailleurs égales, proportionnel à ſon poids & à l'eſpace qu'on lui fait parcourir, & par conſéquent le prix du tranſport total devant être proportionnel à la ſomme des produits des molécules multipliées chacune par l'eſpace parcouru, il s'enſuit que le déblai & le remblai étant donnés de figure & de poſition, il n'eſt pas indifférent que telle molécule du déblai ſoit tranſportée dans tel ou tel autre endroit du remblai, mais qu'il y a une certaine diſtribution à faire des molécules du premier dans le ſecond, d'après laquelle la ſomme de ces produits ſera la moindre poſſible, & le prix du tranſport total ſera un *minimum*.

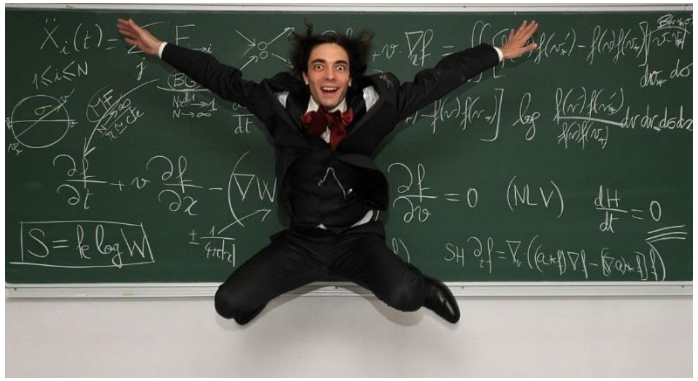
C'eſt la ſolution de cette queſtion que je me propoſe de donner ici. Je diviſerai ce Mémoire en deux parties, dans la première je ſuppoſerai que les déblais & les remblais ſont des aires contenues dans un même plan : dans le ſecond, je ſuppoſerai que ce ſont des volumes.

PREMIÈRE PARTIE.

Du tranſport des aires planes ſur des aires compriſes dans un même plan.

I.

QUELLE que ſoit la route que doive ſuivre une molécule



Cédric Villani: French Mathematician

SECTION COLLOQUIUM 2019

Quantile Combination

Step 1 : joint law

Optimal transport theory main result (applied to or problem)

- Assumption :
  c(x,y) = c(x - y)

- Then :

  1) The optimal dependence structure that reach the previous minimization problem is given by π^*, with joint cdf function $F_{\pi*}(x, y) = \min(F_{\mu}(x), F_{\nu}(y))$,

  2) and the optimal total desagreement is given by: $V_c = \int_0^1 c(q_{\mu}(u) - q_{\nu}(u))du$

=> The optimal dependence structure π^* is **comonotone** structure

Quantile Combination

Step 1 : joint law

Conclusion / remark

1) The comonotone dependency structure minimizes the disagreement between the two models.

2) Quantile appear naturally within « transportation theory formulation »

Questions that remain :

How to choose the weights ?

Quantile Combination

Step 2 : Weights computation

“The weights are chosen in order to minimize the variance of the quantile estimator being constructed”

Quantile combination estimator

Let assume that models corresponds to iid samples (X_1,….X_n) and (Y_1,….Y_m) with μ and ν distribution respectively

Let $X_{(1,n)} < X_{(2,n)} \dots . < X_{(n,n)}$ and $Y_{(1,n)} < Y_{(2,m)} \dots . < Y_{(m,m)}$ « statistic order » associated to X and Y samples

Let u ∈ [0,1] and

$$\widehat{q_\lambda(u)} := \lambda \times X_{(k,n)} + (1-\lambda) \times Y_{(l,m)}$$

With
k = $\lfloor n \times u \rfloor$ +1 and l = $\lfloor m \times u \rfloor$ +1

Quantile Combination

Step 2 : Weights computation

$$\widehat{q_\lambda(u)} := \lambda \times X_{(k,n)} + (1-\lambda) \times Y_{(l,m)}$$

We choose λ in order to minimise $\widehat{q_\lambda}$ variance

Computation of $\widehat{q_\lambda}$ variance leads to

$$\lambda^*{}_u = \max(\min\left(\frac{a}{a+b}, 1\right), 0)$$

With

a = var$(q_\nu(Z_{l,m})) - cov(q_\mu(Z_{k,n}), q_\nu(Z_{l,m}))$
b = var$(q_\mu(Z_{k,n})) - cov(q_\mu(Z_{k,n}), q_\nu(Z_{l,m}))$

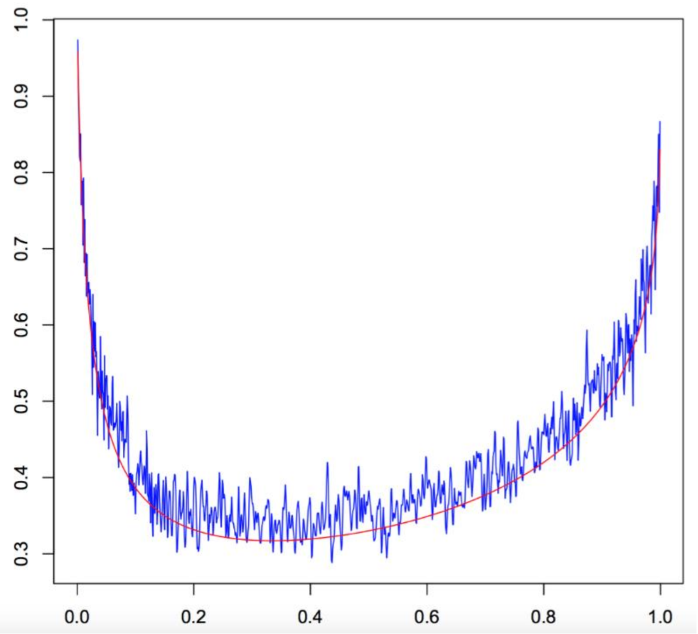
Where $Z_{h,i}$ is a random variable which follows a Beta(h, i – h + 1) distribution

Weights choice

Numerical example

We suppose: X ~ Log-normal and Y ~ Weibull distributions

a and b are estimated throw numerical simulations of the Beta distribution and corresponding X and Y quantile



Limits weights can be computed if X and Y distribution are known…

Application to the modeling of natural disasters

Types of models	Avantages	Drawbacks
Historical models (Statistical approach)	▪ Consistent with observed losses ▪ Easy to model	▪ Does not take into account events that have not occurred ▪ Does not take into account the evolutions of the exposure of a the portfolio ▪ Sensitivity to the addition of a damaged year
Physical models (Exposure approach)	▪ Takes into account extreme events that have not occurred ▪ Allows modeling of all hazards ▪ Takes into account the evolution of the exposure of the portfolio	▪ Instability of modeling results (model change) ▪ Need a lot of modeling assumptions ▪ Not always consistent with observations

Application to the modeling of natural disasters

Consider that we have the opinion of two experts (two models) for a given natural disaster event:

- The first is based on historical events: the historical expert
- The second on physical model: the exposure expert

Both experts agree on the "seriousness" of the event : they agree on the return period corresponding to the event (comonotone hypothesis).

However, they do not evaluate losses related to the event in the same way (they have different models).

Rational hypothesis: When the event is a frequent events, we give more credit to the historical expert and vice versa for extreme events.

Application to the modeling of natural disasters

Mathematical translation: quantile combination of the 2 models

For each return period « u » compute combined quantile:

$$\boldsymbol{q_{combined}(u) = (1 - \lambda_u)} \times \boldsymbol{q_{histo}(u)} + \boldsymbol{\lambda_u} \times \boldsymbol{q_{expo}}(\mathrm{u})$$

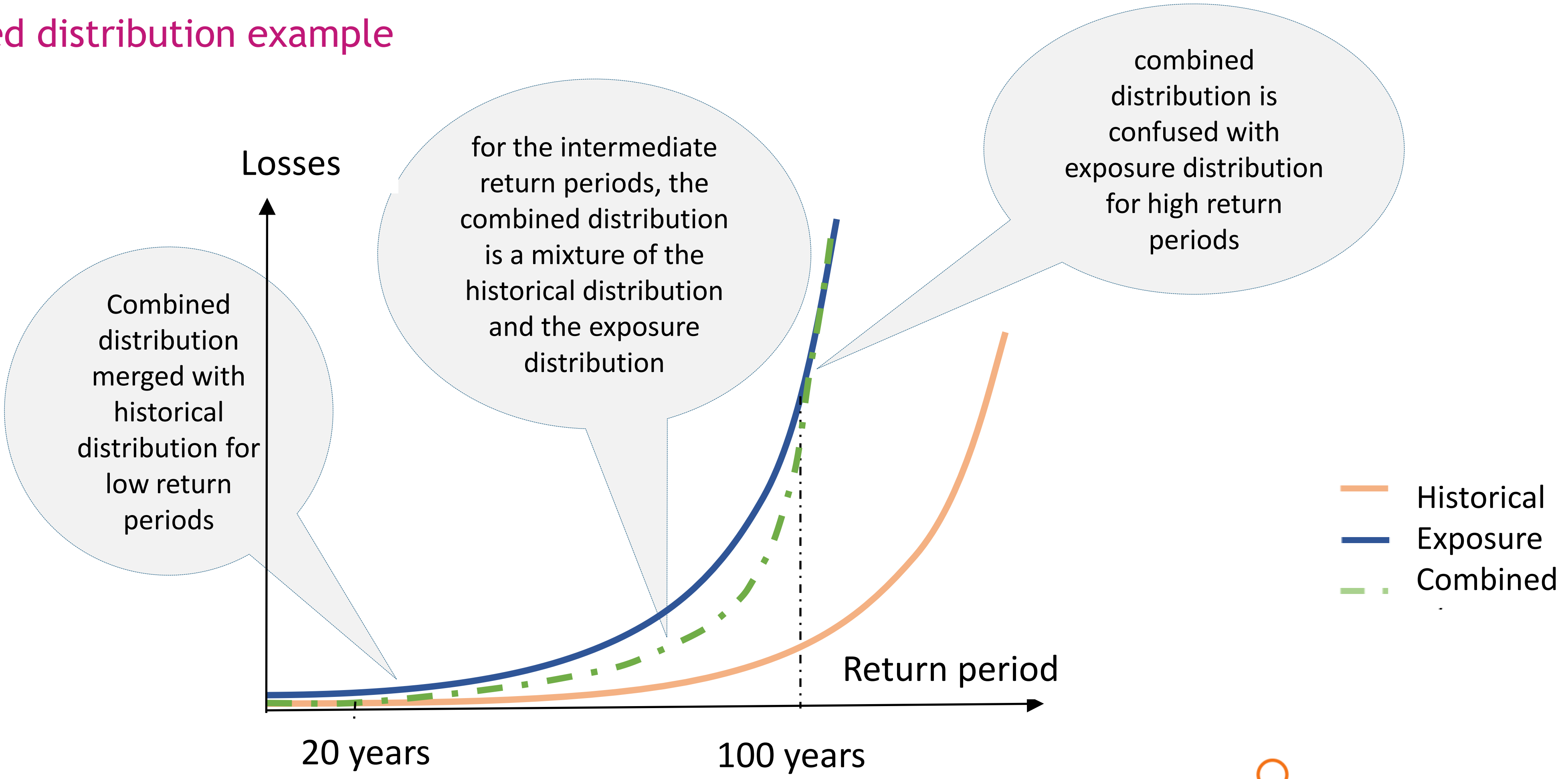
$\boldsymbol{\lambda_u}$ represents the « credibility » accorded to the « exposure » model for return period u

Typically: $\boldsymbol{\lambda_u}$ is close to 1 for high return period (rare events) and conversely, $\boldsymbol{\lambda_u}$ is close de 0 for low return period (frequent events)

Remark: comonotonicity correspond to the fact that the same return period « u » is taken into account for historical and exposure models

Application to the modeling of natural disasters

Combined distribution example



Application to the modeling of natural disasters

	Return period	Avantages
Combined model	Low return period	Consistent with observed losses
	Hight return period	▪ Takes into account extreme events that have not occurred ▪ Allows modeling of all hazards ▪ Takes into account the evolution of the exposure of the portfolio

Thank you for your attention !