Local predictive distributions for modeling risk under parameter and model uncertainty

Léonard Vincent

University of Lausanne

SAV Bahnhofskolloquium, 7 March 2022

Overview

In this presentation

- Risk ... aleatoric/epistemic/model/parameter uncertainties
- Ex-ante and ex-post solvency probabilities
- Estimated and posterior predictive distributions
- Transformed location-scale risk
- Local posterior predictive distribution

Introduction

Quantitative risk ...

- e.g. insurance loss (stock price, earthquake magnitude, etc.)
- Modelled as a random variable ... e.g. $X \sim F_X(x)$
- $F_X(x) = P(X \leq x)$
- $F_X(x) \Rightarrow$ measure of risk ... e.g. quantiles, std, etc.

In practice ... $F_X(x)$ is unknown

- Data $\mathbf{d} = (d_1, ..., d_n) \Rightarrow \hat{F}_X(x; \mathbf{d})$ ... estimated distribution (best fit)

Assume $X \sim \hat{F}(x; \mathbf{d})$? ... Ignore risk that $\hat{F}(x; \mathbf{d}) \neq F(x)$?

Aleatoric and epistemic uncertainties

Two types of uncertainty

- $X \sim F(x)$ ... aleatoric uncertainty
- $\hat{F}(x; \mathbf{d}) \neq F(x)$ ... epistemic uncertainty

Epistemic uncertainty shall be considered ... e.g. risk trade-off

- 2 independent risks X and Y
- Data $\mathbf{d}_X$ and $\mathbf{d}_Y$
- Assume $\hat{F}_X(x; \mathbf{d}_X) = \hat{F}_Y(x; \mathbf{d}_Y)$ ... allocate same resources ?
- Size of $\mathbf{d}_X$ and $\mathbf{d}_Y$?

See e.g. Gerrard and Tsanakas (2011), Fröhlich and Weng (2015), Pitera and Schmidt (2018)

Model and parametric uncertainties

Epistemic uncertainty under parametric approach

- Assume $F_X(x) = F_X(x; \boldsymbol{\theta})$ ... parameter $\boldsymbol{\theta} = (\theta_1, \theta_2, ...)$
- $\mathbf{d} \Rightarrow \hat{\boldsymbol{\theta}}$ ... estimated parameter (e.g. MLE)
- $F_X(x; \hat{\boldsymbol{\theta}})$ ... estimated distribution

Source for $F_X(x; \hat{\boldsymbol{\theta}}) \neq F_X(x)$? (epistemic uncertainty)

- $F_X(x; \boldsymbol{\theta}) \neq F_X(x)$ ... wrong model (model uncertainty)
- $\hat{\boldsymbol{\theta}} \neq \boldsymbol{\theta}$ ... wrong parameter (parameter uncertainty)

Solvency probability

Assume only parameter uncertainty ... e.g.

- $X \sim F_X(x; \boldsymbol{\theta})$ ... $\boldsymbol{\theta}$ unknown
- Target solvency probability β ... e.g. $\beta = 0.995$
- Theoretical capital $Q_X(\beta; \boldsymbol{\theta}) = F_X^{-1}(\beta; \boldsymbol{\theta})$ ... β-quantile

With data

- $\mathbf{D} = (X_1, ..., X_n)$ ... e.g. iid past realizations of X
- $\mathbf{d} = (x_1, ..., x_n) \Rightarrow \hat{\boldsymbol{\theta}}$ ... MLE

Hold capital $Q_X(\beta; \hat{\boldsymbol{\theta}})$?

Ex-ante and ex-post solvency probabilities

Hold capital $Q_X(\beta; \hat{\boldsymbol{\theta}})$?

- Ex-post solvency probability ... $P(X \leq Q_X(\beta; \hat{\boldsymbol{\theta}}); \boldsymbol{\theta}) \neq \beta$
- Ex-ante solvency probability ... $P(X \leq Q_X(\beta; \hat{\boldsymbol{\Theta}}); \boldsymbol{\theta}) \leq \beta$

Why ? $Q_X(\beta; \hat{\boldsymbol{\theta}})$ calculated with $F_X(x; \hat{\boldsymbol{\theta}})$

$\Rightarrow$ As if $X \sim F_X(x; \hat{\boldsymbol{\theta}})$

$\Rightarrow$ Ignores epistemic (parameter) uncertainty

Pareto example

$X \sim$ Pareto with unknown scale and shape parameters ... For $n \geq 2$

$$P(X \leq Q_X(\beta; \hat{\boldsymbol{\Theta}}); \boldsymbol{\theta}) = 1 - \left(\frac{n}{n+1}\right)\left(1 - n^{-1}\log(1-\beta)\right)^{-n+1} \leq \beta$$

e.g. $\beta = 99.5\%$ ... Then $P(X \leq Q_X(\beta; \hat{\boldsymbol{\Theta}}); \boldsymbol{\theta}) =$

- 98.0% ... for $n = 10$
- 98.9% ... for $n = 20$
- 99.4% ... for $n = 100$

Bayesian setup

Ω ... parameter space

- Distribution family $\mathcal{F} = \{F_X(x; \boldsymbol{\theta}), \boldsymbol{\theta} \in \Omega\}$
- $\mathcal{F}$ ... possible descriptions of X's aleatoric uncertainty
- $F_X(x; \hat{\boldsymbol{\theta}})$ ... only one of them

Bayesian approach

- $\boldsymbol{\theta} \Rightarrow \boldsymbol{\Theta}$ ... random parameter
- $\boldsymbol{\Theta} \sim f_{\boldsymbol{\Theta}}(\boldsymbol{\theta})$ ... prior (epistemic uncertainty)

Posterior predictive distribution

Prior $f_{\boldsymbol{\Theta}}(\boldsymbol{\theta}) \Rightarrow$ Posterior

$$f_{\boldsymbol{\Theta}|\mathbf{D}}(\boldsymbol{\theta}|\mathbf{d}) = \frac{f_{\mathbf{D}}(\mathbf{d};\boldsymbol{\theta})f_{\boldsymbol{\Theta}}(\boldsymbol{\theta})}{\int_{\Omega} f_{\mathbf{D}}(\mathbf{d};\boldsymbol{\theta}')f_{\boldsymbol{\Theta}}(\boldsymbol{\theta}')\,\mathrm{d}\boldsymbol{\theta}'}$$

$\Rightarrow$ Posterior predictive distribution (ppd)

$$F_{X|\mathbf{D}}(x|\mathbf{d}) = \int_{\Omega} F_X(x;\boldsymbol{\theta})f_{\boldsymbol{\Theta}|\mathbf{D}}(\boldsymbol{\theta}|\mathbf{d})\,\mathrm{d}\boldsymbol{\theta}$$

- Previously: $F_X(x;\hat{\boldsymbol{\theta}})$ ... "modal" distribution
- Now: $F_{X|\mathbf{D}}(x|\mathbf{d})$ ... "average" distribution

Choice of $f_{\boldsymbol{\Theta}}(\boldsymbol{\theta})$?

Choice of prior

Choice of $f_{\Theta}(\boldsymbol{\theta})$? With "rule" (subjective Bayes) ... e.g.

- $f_{\Theta}(\boldsymbol{\theta}) > 0$ for all $\boldsymbol{\theta} \in \Omega$ ... allows for asymptotic convergence
- Conjugate prior ... ease of computation
- etc.

Only "rule" (objective Bayes) ... e.g.

- Invariance w.r.t. re-parametrization ... Jeffreys prior
- Adequacy with frequentist criterion ... Matching prior
- etc.

Matching prior

$F_{X|\mathbf{D}}(x|\mathbf{d})$ ... depends on $f_{\boldsymbol{\Theta}}(\boldsymbol{\theta})$

- $Q_{X|\mathbf{D}}(\beta|\mathbf{d}) = F^{-1}_{X|\mathbf{D}}(\beta|\mathbf{d})$ ... Predictive β-quantile

Then $f_{\boldsymbol{\Theta}}(\boldsymbol{\theta})$ is a matching prior if

$$P(X \leq Q_{X|\mathbf{D}}(\beta|\mathbf{D}); \boldsymbol{\theta}) = \beta, \text{ for all } \boldsymbol{\theta} \in \Omega, \beta \in [0,1]$$

... correct ex-ante solvency probability

- Region $(-\infty, Q_{X|\mathbf{D}}(\beta|\mathbf{d})]$
- Bayesian coverage probability $F_{X|\mathbf{D}}(Q_{X|\mathbf{D}}(\beta|\mathbf{d})|\mathbf{d}) = \beta$
- Frequentist coverage probability $P(X \leq Q_{X|\mathbf{D}}(\beta|\mathbf{D}); \boldsymbol{\theta})$

Transformed location-scale family

$\mathcal{F}^{tls}$ … transformed location-scale (tls) family

- $X = g(\theta_1 + \theta_2 Z)$
- $Z \sim F_Z(x)$ … known distribution
- $g(x)$ … strictly increasing function
- $\boldsymbol{\theta} = (\theta_1, \theta_2) \in \Omega = \mathbb{R} \times \mathbb{R}_+$

Examples

- $Z \sim \text{Exp}(1)$ … $X = e^{\theta_1 + \theta_2 Z} \sim \text{Pareto}(e^{\theta_1}, \theta_2^{-1})$
- $Z \sim \text{Normal}(0, 1)$ … $X = e^{\theta_1 + \theta_2 Z} \sim \text{LogNormal}(\theta_1, \theta_2^2)$
- Weibull, Normal, Uniform, etc.

Transformed location-scale family

For $\mathcal{F}^{tls}$... with $n \geq 2$

- $f_{\boldsymbol{\Theta}}(\boldsymbol{\theta}) \propto \theta_2^{-1}$, for all $\boldsymbol{\theta} \in \Omega$ ... matching prior
- $\int_\Omega \theta_2^{-1} \, \mathrm{d}\boldsymbol{\theta} = \infty$ ... improper pdf

Recall ... Bayes formula $f_{\boldsymbol{\Theta}|\mathbf{D}}(\boldsymbol{\theta}|\mathbf{d}) = \frac{f_{\mathbf{D}}(\mathbf{d};\boldsymbol{\theta}) f_{\boldsymbol{\Theta}}(\boldsymbol{\theta})}{\int_\Omega f_{\mathbf{D}}(\mathbf{d};\boldsymbol{\theta}') f_{\boldsymbol{\Theta}}(\boldsymbol{\theta}') \, \mathrm{d}\boldsymbol{\theta}'}$

In general for continuous X and $P(X_1 = ... = X_n) = 0$

- $f_{\boldsymbol{\Theta}|\mathbf{D}}(\boldsymbol{\theta}|\mathbf{d})$ and $F_{X|\mathbf{D}}(x|\mathbf{d})$ ... proper a.s.

In which case $\Rightarrow P(X \leq Q_{X|\mathbf{D}}(\beta|\mathbf{D}); \boldsymbol{\theta}) = \beta$, for all $\boldsymbol{\theta} \in \Omega, \beta \in [0,1]$

Extends Severini et al. (2002), Gerrard and Tsanakas (2011)

Pareto example

- $\mathbf{D} = (X_1, ..., X_n) \overset{iid}{\sim}$ Pareto ... $n \geq 2$
- $\mathbf{d} = (x_1, ..., x_n)$
- $t_1 = \min(\mathbf{d})$
- $t_2 = n / \sum_{i=1}^{n} \ln(x_i / t_1)$

Then ppd

$$F_{X|\mathbf{D}}(x|\mathbf{d}) = \begin{cases} F_{X|\mathbf{D}}(x|\mathbf{d}) > 0 & , 0 < x < t_1 \\ 1 - \left(\frac{n}{n+1}\right)\left(1 + \ln\left[\left(\frac{t_1}{x}\right)^{t_2}\right]\right)^{-n+1} & , x \geq t_1 \end{cases}$$

Compare with estimated distribution

$$F_X(x; \hat{\boldsymbol{\theta}}) = \begin{cases} 0 & , x < t_1 \\ 1 - \left(\frac{t_1}{x}\right)^{t_2} & , x \geq t_1 \end{cases}$$

With model uncertainty

Model uncertainty

- $\mathcal{F} = \{F_X(x;\boldsymbol{\theta}), \boldsymbol{\theta} \in \Omega\}$ ... rarely perfectly known
- Bayes with several models ... loose matching property

Parametric model "locally" appropriate ... e.g.

- Pareto or Log-Normal tail
- (Log-)Normal body of distribution
- etc.

$\Rightarrow$ "local" posterior predictive distribution

Local posterior predictive distribution

Setup

- $\mathbf{D} = (X_1, ..., X_k) \overset{iid}{\sim} F_X(x)$
- $\mathbf{d} = (x_1, ..., x_k)$
- $\mathcal{F}^{tls}$ ... appropriate for $x \in (b_\ell, b_u)$
- $x_{i:k}$ ... i^{th} greatest value in $\mathbf{d}$ (order statistic)

$$\underbrace{x_{1:k} < ... < x_{m:k}}_{m \geq 0} < b_\ell < \underbrace{x_{m+1:k} < ... < x_{m+n:k}}_{n \geq 0} < b_u < \underbrace{x_{m+n+1:k} < ... < x_{k:k}}_{k-m-n \geq 0}$$

- $\tilde{\mathbf{d}} = (x_{m+1:k}, ..., x_{m+n:k})$
- $\tilde{\mathbf{D}} = (X_{M+1:k}, ..., X_{M+N:k})$

Local posterior predictive distribution

In general ... for

- $f_{\mathbf{\Theta}}(\boldsymbol{\theta}) \propto \theta_2^{-1}$, for all $\boldsymbol{\theta} \in \Omega$
- X continuous in (b_ℓ, b_u)
- $P(X_1 = ... = X_k) = 0$

Then

$$F_{X|\tilde{\mathbf{D}}}(x|\tilde{\mathbf{d}}) = \int_\Omega F_X(x;\boldsymbol{\theta}) \frac{f_{\tilde{\mathbf{D}}}(\tilde{\mathbf{d}};\boldsymbol{\theta})\theta_2^{-1}}{\int_\Omega f_{\tilde{\mathbf{D}}}(\tilde{\mathbf{d}};\boldsymbol{\theta}')\theta_2'^{-1}\,\mathrm{d}\boldsymbol{\theta}'}\,\mathrm{d}\boldsymbol{\theta}$$

proper a.s. ... given $n \geq 2$.

Remember

- $F_{X|\tilde{\mathbf{D}}}(x|\tilde{\mathbf{d}})$ ... relevant for $x \in (b_\ell, b_u)$

Local posterior predictive distribution

Correct ex-ante solvency probability ? Requires $\tilde{\mathbf{D}}$ such that

(*i*) Proper $F_{X|\tilde{\mathbf{D}}}(x|\tilde{\mathbf{d}})$ ... $N \geq 2$

(*ii*) $Q_{X|\tilde{\mathbf{D}}}(\beta|\tilde{\mathbf{d}}) \in (b_\ell, b_u)$

For a given $\beta \in [0, 1]$ ... Define

- $\mathcal{D} = \left\{ \tilde{\mathbf{d}} : (i) \text{ and } (ii) \text{ are satisfied} \right\}$
- $\tilde{\mathbf{D}}^* = (\tilde{\mathbf{D}} | \tilde{\mathbf{D}} \in \mathcal{D})$

Correct ex-ante "conditional" solvency probability

$$P(X \leq Q_{X|\tilde{\mathbf{D}}}(\beta|\tilde{\mathbf{D}}^*); \boldsymbol{\theta}) = \beta, \text{ for all } \boldsymbol{\theta} \in \Omega$$

Pareto example

- $\mathbf{D} = (X_1, ..., X_k) \overset{iid}{\sim} F_X(x)$
- Pareto ... appropriate for $x \in (b_\ell, \infty)$
- $\tilde{\mathbf{d}} = (x_{k-n:k}, ..., x_{k:k})$ ... $n \geq 2$
- $t_1 = x_{k-n:k}$ and $t_2 = \sum_{i=k-n}^{k} \ln(x_{i:k}/t_1)$

Local ppd

$$F_{X|\tilde{\mathbf{D}}}(x|\tilde{\mathbf{d}}) = 1 - \left(\frac{k-m}{k+1}\right)\left(1 + \ln\left[\left(\frac{t_1}{x}\right)^{t_2}\right]\right)^{-n+1}, \text{ for } x > b_\ell$$

Conclusion

- Aleatoric/epistemic/model/parameter uncertainties
- Ex-ante/ex-post solvency probabilities
- Posterior predictive distribution for tls
- Local posterior predictive distribution

Paper soon available... Interested ? Contact me

- leonard.vincent@unil.ch

References I

Fröhlich, A. and Weng, A. (2015). Modelling parameter uncertainty for risk capital calculation. *European Actuarial Journal*, 5(1):79–112.

Gerrard, R. and Tsanakas, A. (2011). Failure probability under parameter uncertainty. *Risk Analysis: An International Journal*, 31(5):727–744.

Pitera, M. and Schmidt, T. (2018). Unbiased estimation of risk. *Journal of Banking & Finance*, 91:133–145.

Severini, T. A., Mukerjee, R., and Ghosh, M. (2002). On an exact probability matching property of right-invariant priors. *Biometrika*, 89(4):952–957.