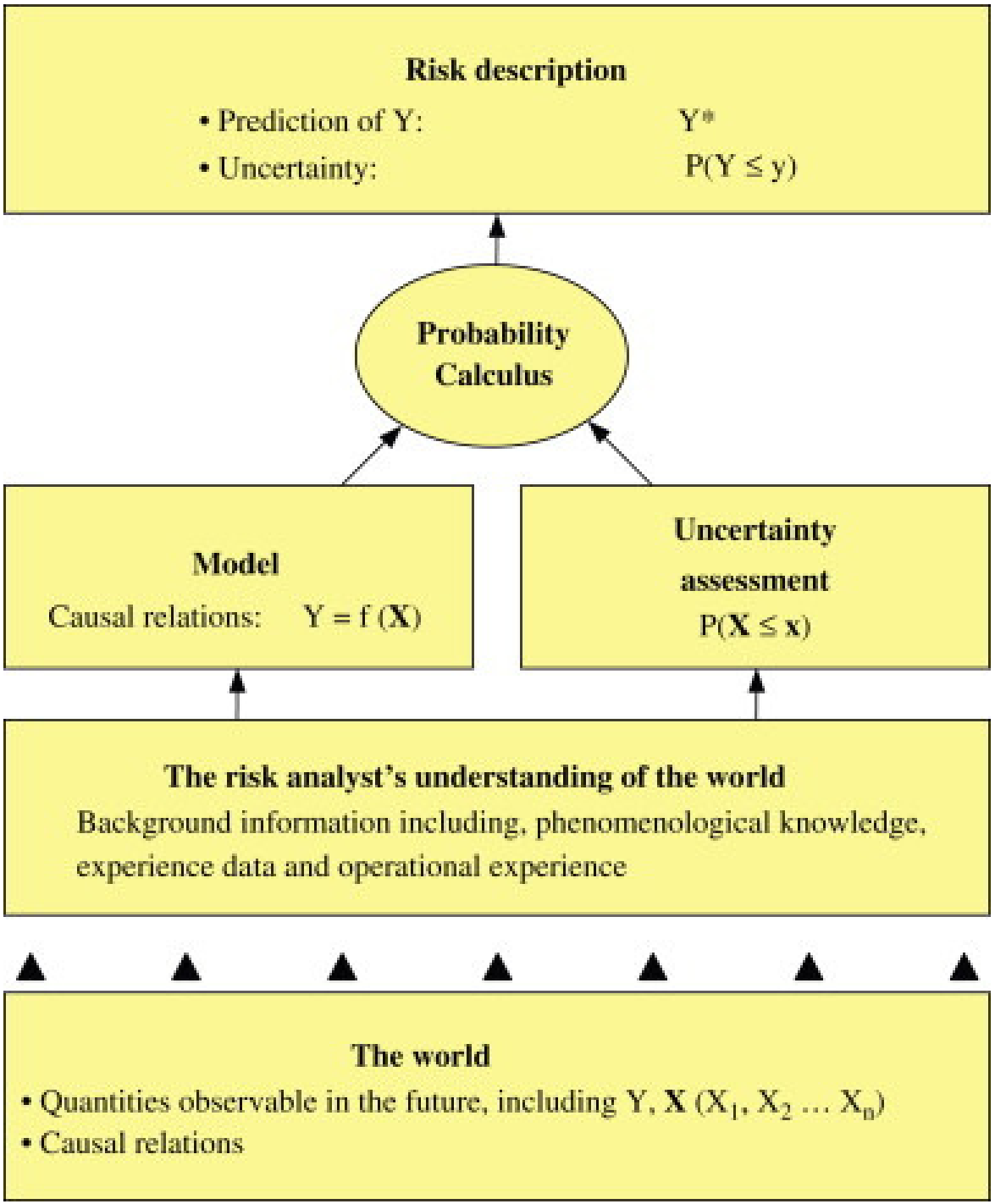
Uncertainty in QSAR predictions for probabilistic risk assessment

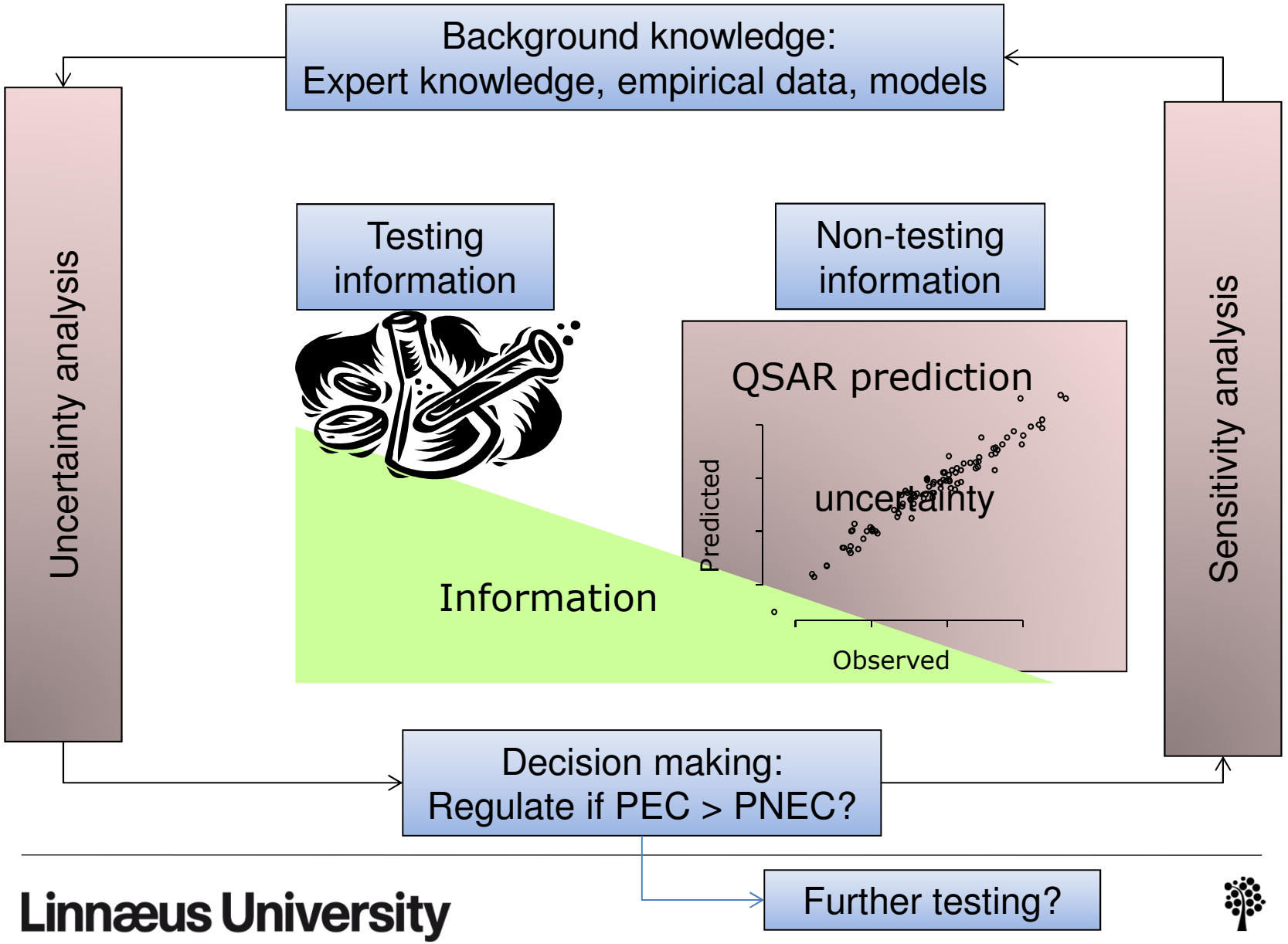
Ullrika Sahlin PhD

Linnæus University

Why uncertainty is important

Decision makers need to be aware of the degree of uncertianty associated with the results of the evaluation of the available scientific knowlege


Risk description
• Prediction of Y: Y*
• Uncertainty: P(Y ≤ y)
Probability Calculus
Model
Causal relations: Y = f (X)
Uncertainty assessment
P(X ≤ x)
The risk analyst's understanding of the world
Background information including, phenomenological knowledge, experience data and operational experience
The world
• Quantities observable in the future, including Y, X (X1, X2 … Xn)
• Causal relations


Background knowledge:
Expert knowledge, empirical data, models
Uncertainty analysis
Sensitivity analysis
Testing information
Non-testing information
QSAR prediction
Predicted
uncertainty
Observed
Information
Decision making:
Regulate if PEC > PNEC?
Further testing?

Uncertainty analysis

Highligth those areas of uncertainty that have the greatest impact on the analysis

Indicate to the regulator the degree of seriousness of the risk under consideration

Probabilistic risk assessment

a.n.a. Quantitative risk assessment
Pate-Cornell's six levels of treatment from 1996

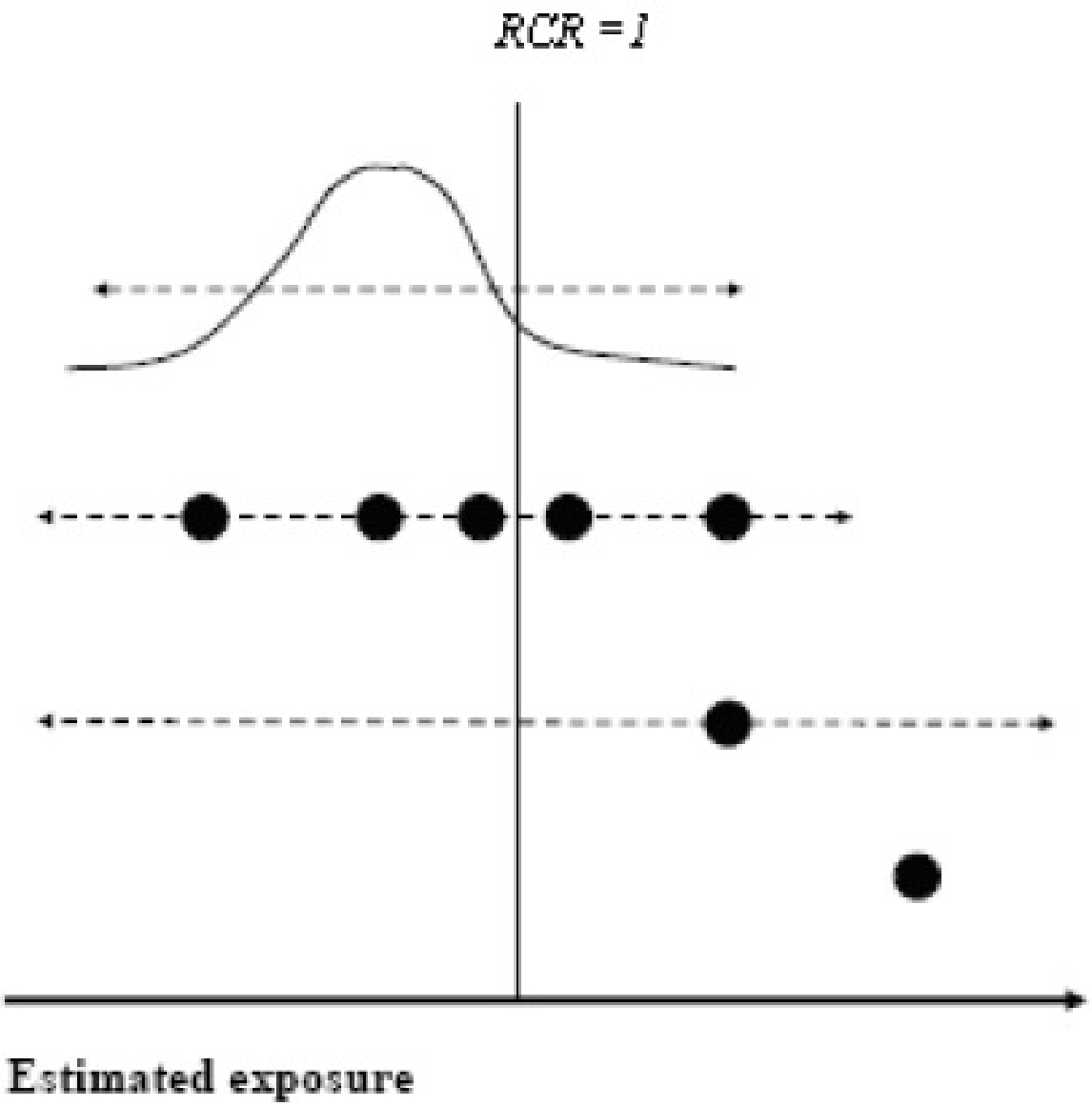
1. Identify hazard
2. Worst case
3. Worst plausible case
4. Best estimates (central values)
5. Probability – single distribution

weight values after how likely they are

6. Probability – several distributions

weight value and allow the weights to be uncertain

REACH – tiered approach

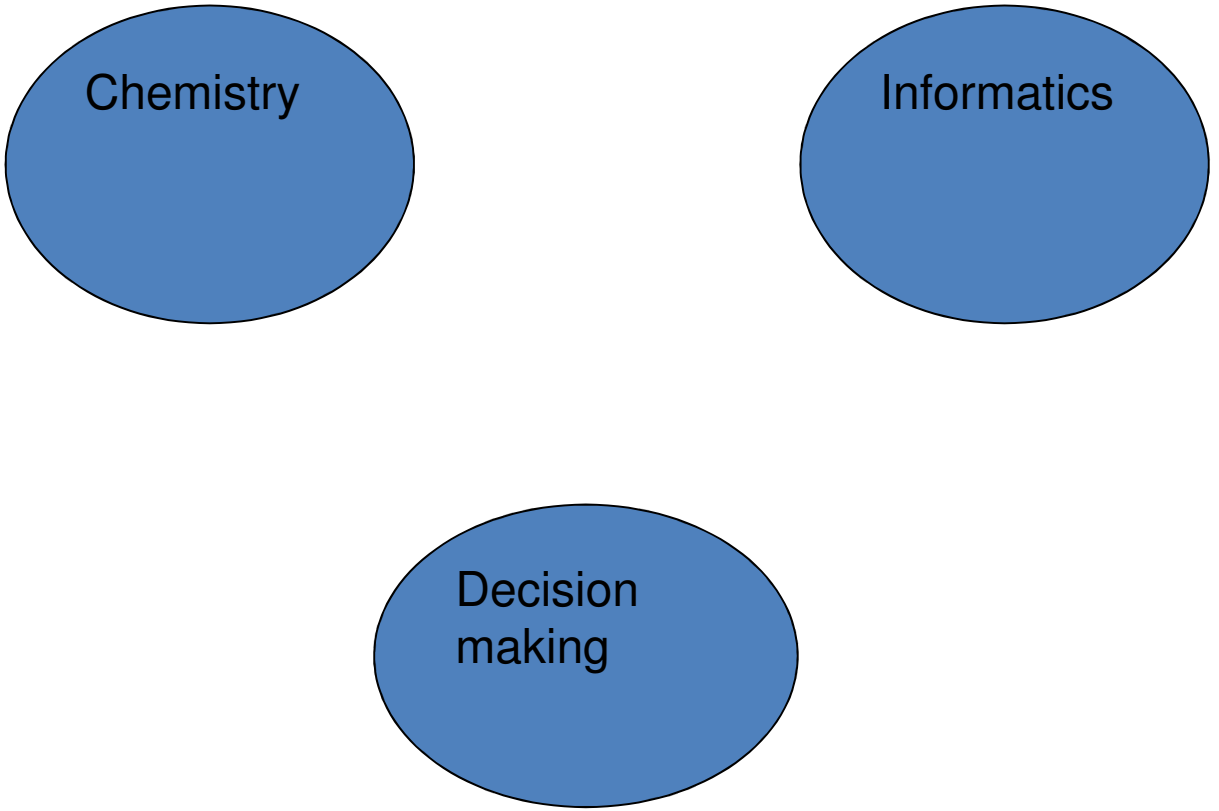


3. Probabilistic – probability distribution for quantified uncertainties, plus indicative range for unquantifiable uncertainties

2. Deterministic – a range of point estimates based on different combinations of assumptions, plus unquantifiable uncertainties

1. Qualitative – refined point estimate plus indicative range for unquantifiable uncertainties

0. Point estimate with conservative assumptions and default values


Chemistry
Informatics
Decision
making

What is a QSAR prediction?

Prediction - a statement about what is thought will happen in the future, often associated with probability distributions

QSAR predict using analogy reasoning

non-parametric: prediction rule

with or without probabilistic model

What is a QSAR prediction?

parametric: mathematical function with parameters

$$Y|x = \beta_0 + \beta_1 x$$
$$y_i = \beta_0 + \beta_1 x_i + e_i$$

Predictive inference is made by assuming a probability model e.g. that all e's are independent and drawn from the same probability distribution such as $N(0, \sigma)$ then if we have enough many data points the prediction will converge to the true value

$$Y|w \sim \beta_0 + \beta_1 w + P(e)$$

or saying that we believe that e, β_0, and β_1 are normally distributed, that σ has a specific distribution and by applying Bayes rule we can derive what we belive the prediction to be

$$Y|w \sim P(\beta_0 + \beta_1 w + e)$$

What is a QSAR prediction? – personalistic and all inclusive view

Query compound: the chemical for which a property/activity is to be predicted

Training data – data set with known values on a quantiative measure of the property/activity

Test/Validation data – same as training data, but not used to build/train the model

Supervised learning algorithm – to train a prediction rule or find estimate a parametric function

QSAR model = Supervised learning algorithm + QSAR data

QSAR prediction = Supervised learning algorithm + QSAR data + descriptors for query compound

+ algorithm to assess uncertainty in prediction

Uncertainty in a QSAR prediction

Two kinds of uncertainty

- Predictive reliability – Qualitative uncertainty
- Predictive distribution / error – Quantitative uncertainty

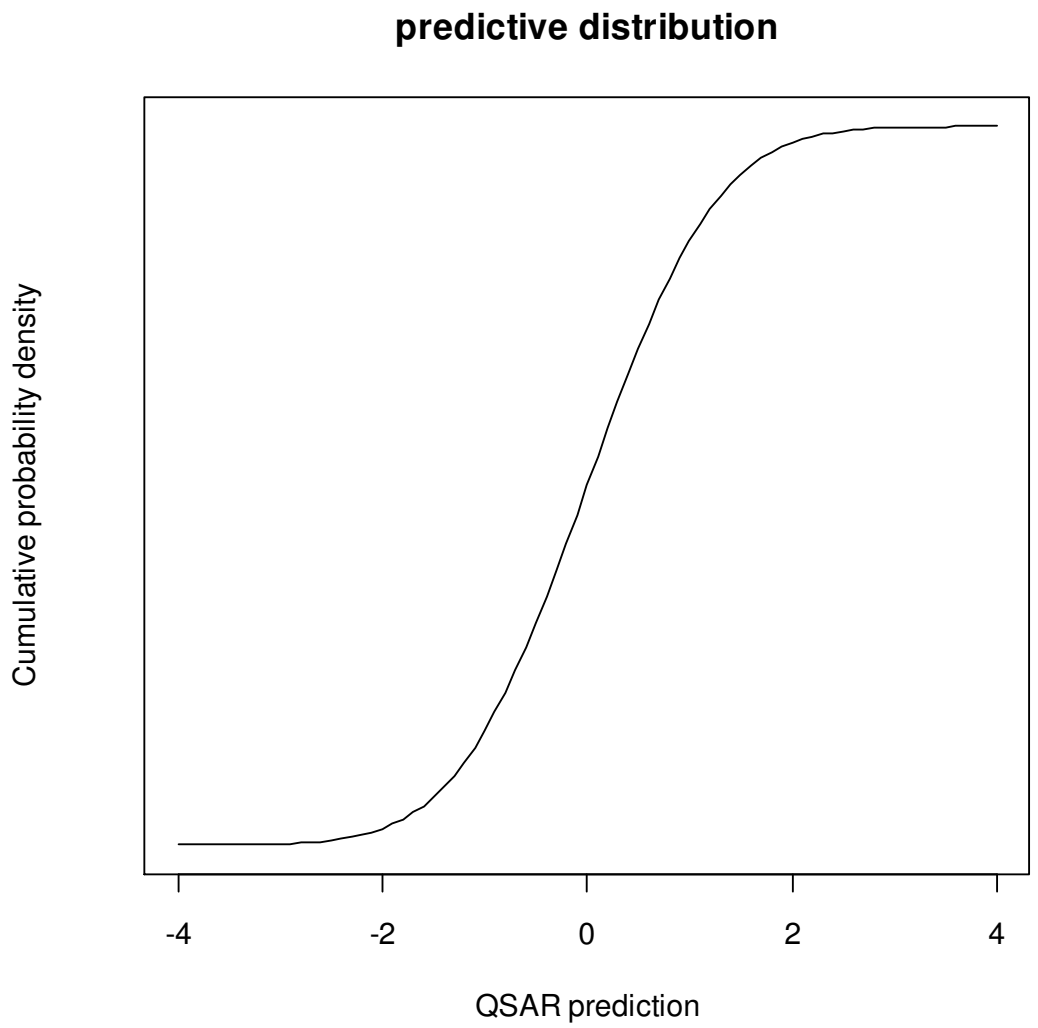
Relation between predictive reliability and predicitve error

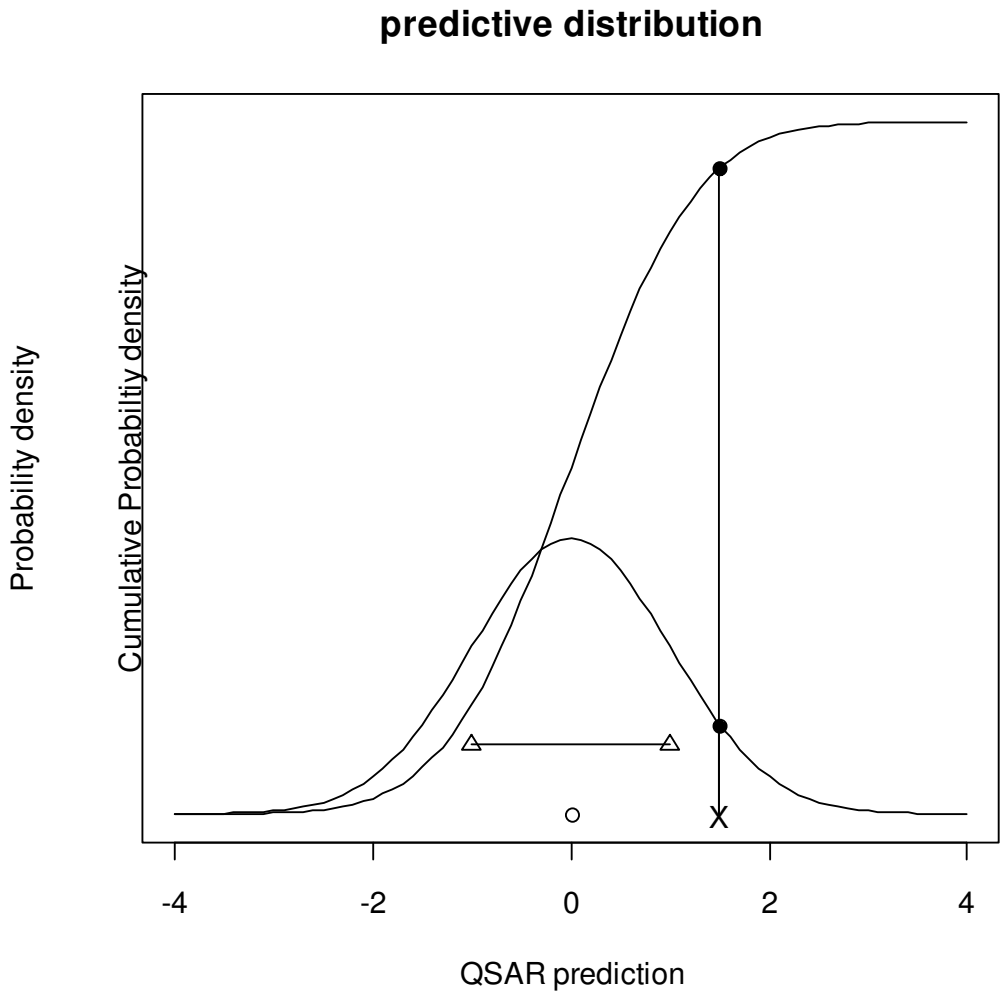
Different QSARs

- Classification models
  - Accuracy, sensitivity and specificity
- Regression models
  - Gaussian-like, symmetric errors
  - Generalized errors (Poisson, Logistic, …)

Need to go from point predictions to

Predictive distribution = our uncertainty in the prediction



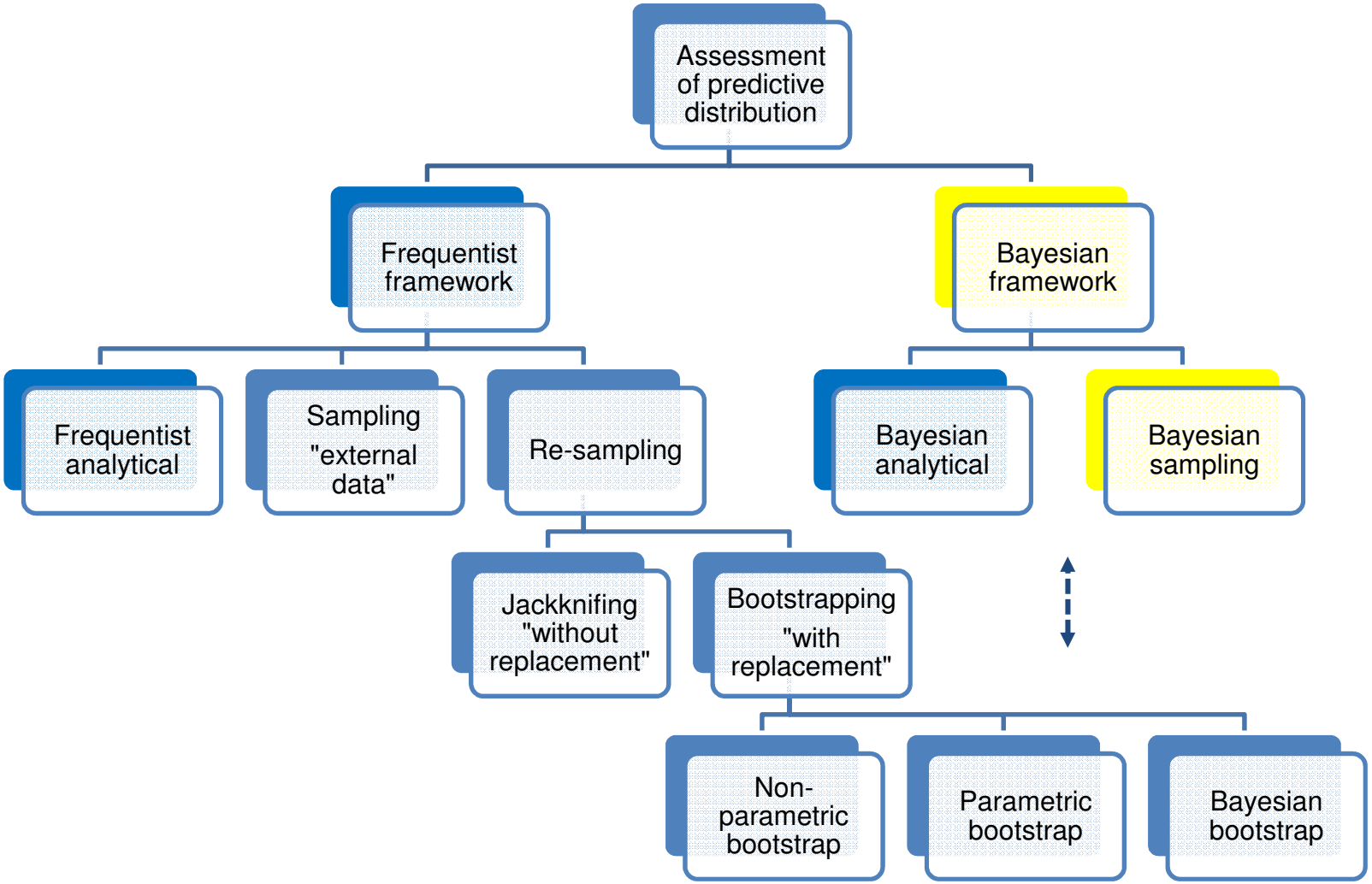

predictive distribution
Probability density
Cumulative Probabiltiy density
-4
-2
0
2
4
QSAR prediction

Predictive reliability

How to characterize and evaluate predictive reliability
How to consider it in hazard or risk assessment?

Related to the Applicability Domain
Clear cut border or smooth – flexibility to the intended application?
Supportive measure as priors of predictive reliability:
- distance to the AD,
- density of the AD,
- spread in ensemble model predictions

Measures for evaluation
- Confidence – ”how much are you willing to bet to trust this model”
- Empirical coverage
- Information measures based on likelihoods – only relative


Assessment of predictive distribution
Frequentist framework
Bayesian framework
Frequentist analytical
Sampling "external data"
Re-sampling
Bayesian analytical
Bayesian sampling
Jackknifing "without replacement"
Bootstrapping "with replacement"
Non-parametric bootstrap
Parametric bootstrap
Bayesian bootstrap

Bayesian model

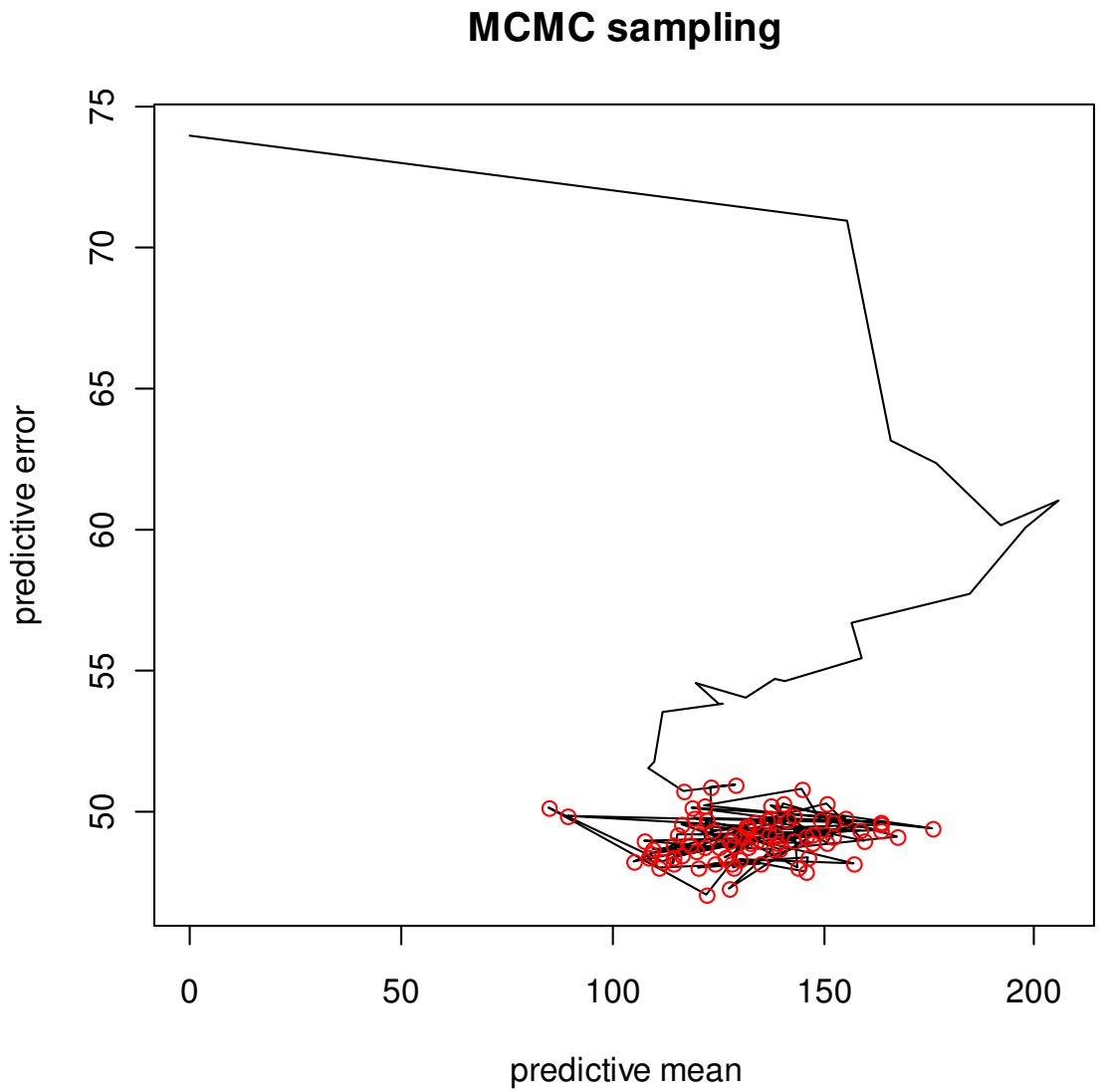
Model is uncertain

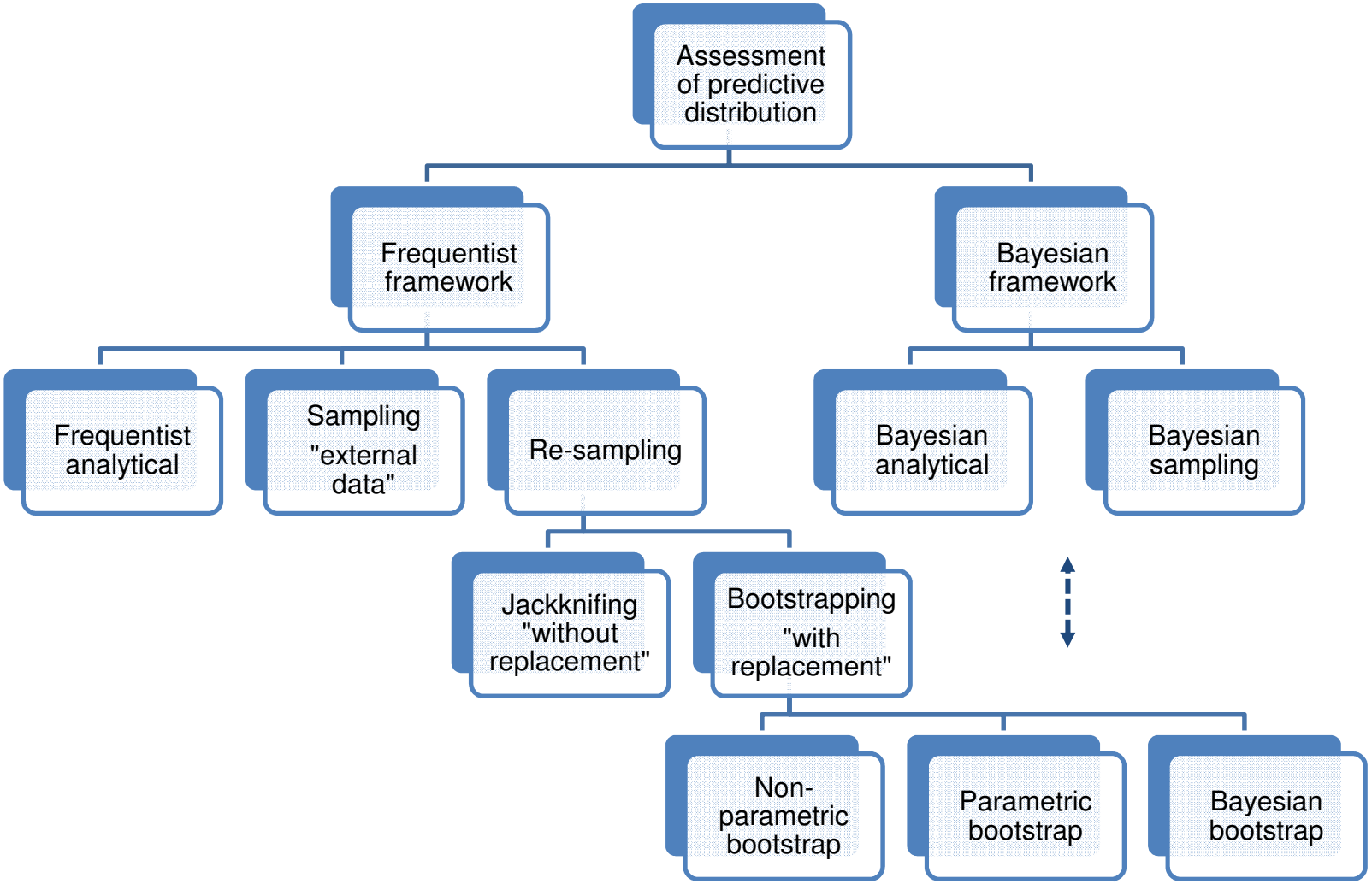
$$Y|w \sim P(\beta_0 + \beta_1 w + e)$$

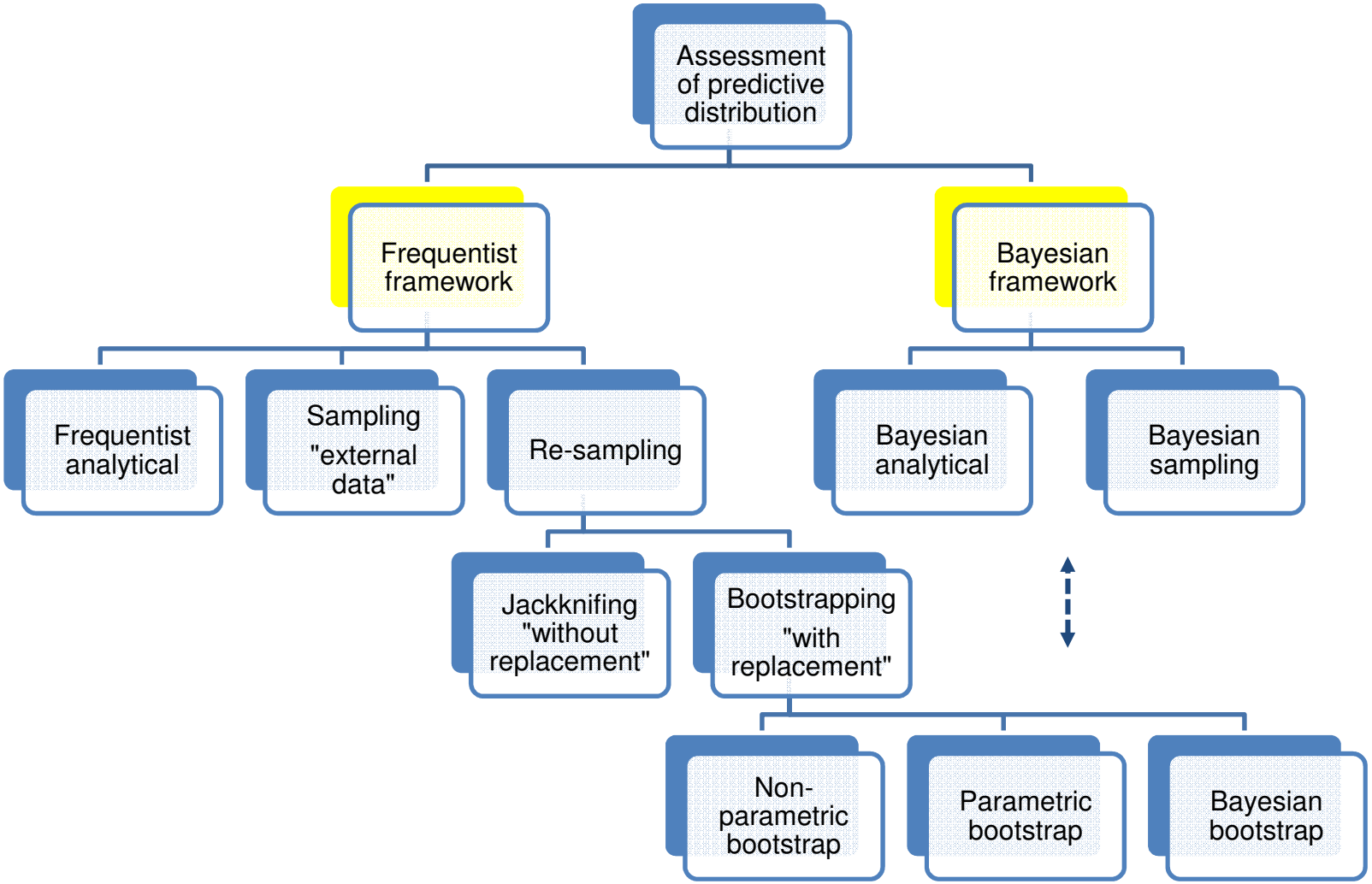
Model is specified by a joint posterior distribution of all the parameters

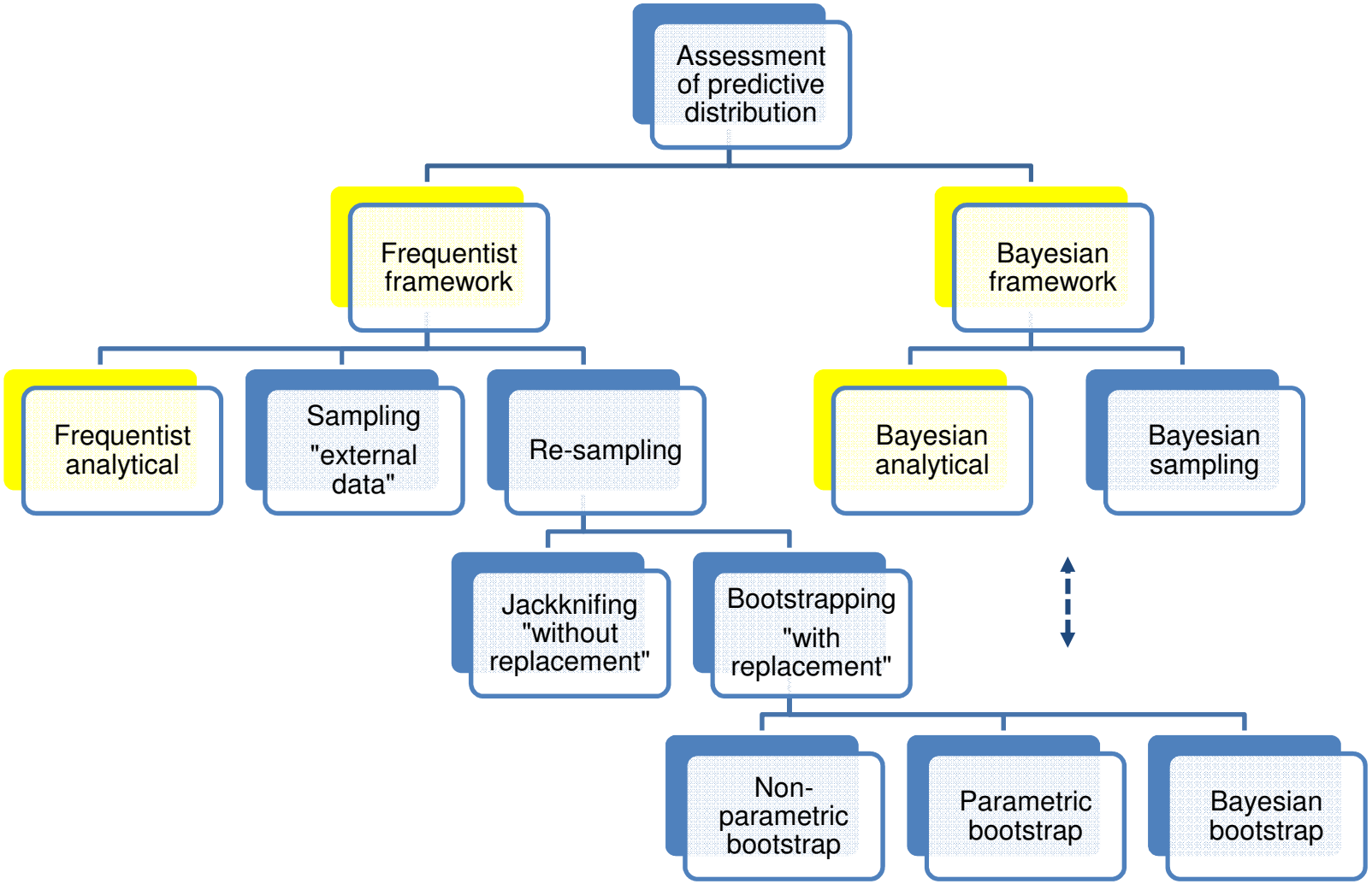
Values are extracted from this posterior by e.g. Markov Chain Monte Carlo sampling

A prediction is given as a sample from the predictive distribution of the query compound


MCMC sampling
predictive error
predictive mean
50
55
60
65
70
75
0
50
100
150
200


Assessment of predictive distribution
Frequentist framework
Bayesian framework
Frequentist analytical
Sampling "external data"
Re-sampling
Bayesian analytical
Bayesian sampling
Jackknifing "without replacement"
Bootstrapping "with replacement"
Non-parametric bootstrap
Parametric bootstrap
Bayesian bootstrap


Assessment of predictive distribution
Frequentist framework
Bayesian framework
Frequentist analytical
Sampling "external data"
Re-sampling
Bayesian analytical
Bayesian sampling
Jackknifing "without replacement"
Bootstrapping "with replacement"
Non-parametric bootstrap
Parametric bootstrap
Bayesian bootstrap


Assessment of predictive distribution
Frequentist framework
Bayesian framework
Frequentist analytical
Sampling "external data"
Re-sampling
Bayesian analytical
Bayesian sampling
Jackknifing "without replacement"
Bootstrapping "with replacement"
Non-parametric bootstrap
Parametric bootstrap
Bayesian bootstrap

Student t - distribution

When: Parametric linear model. Model errors assumed to be independent and identically normally distributed. Fitted by Ordinary Least Square

Information needed

- Predictive mean PRED(Y|W),
- Predictive error SEP(Y|W),
- Number of data points in the training data set (n), and
- Number of descriptors in the linear regression model (k).

The predictive error is estimated as

$$SEP(Y|W)^2 = \sigma^2(1+W^t(X^tX)^{-1}W),$$

where σ^2 is the variance in model errors and $(X^tX)^{-1}$ is the information matrix.

The prediction Y|W was distributed according to its predictive distribution

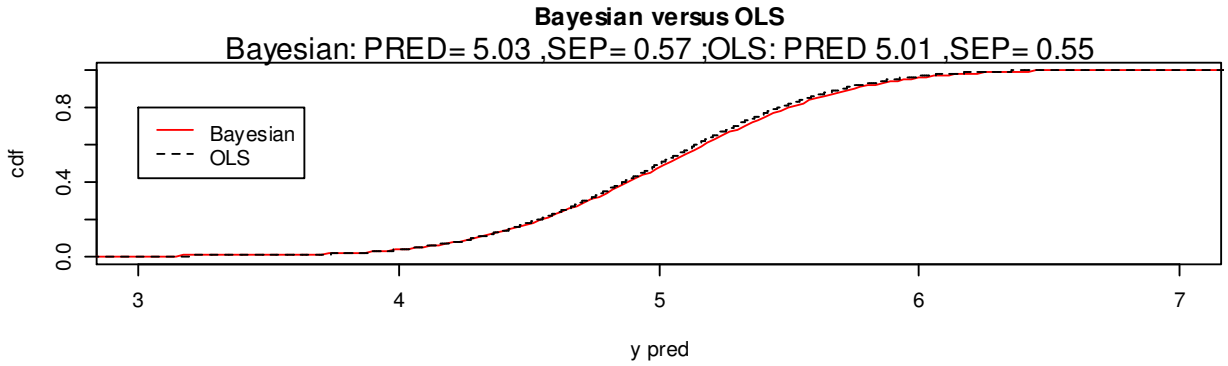
$$Y|W \sim PRED(Y|W)+t_{n-k-1}\ SEP(Y|W)$$

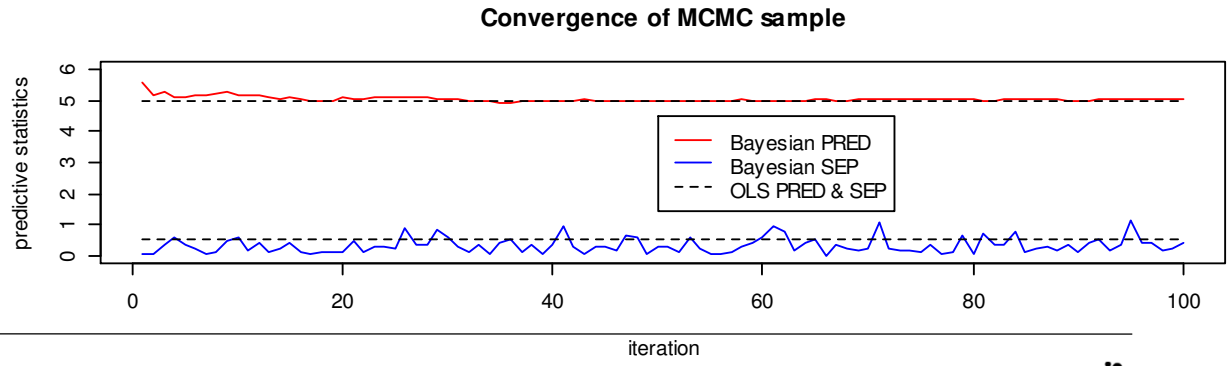
where t_{n-k-1} stands for the t-distribution with n – k –1 degrees of freedom.

Sample from Bayesian with non-informative priors

≈

Frequentist analytical

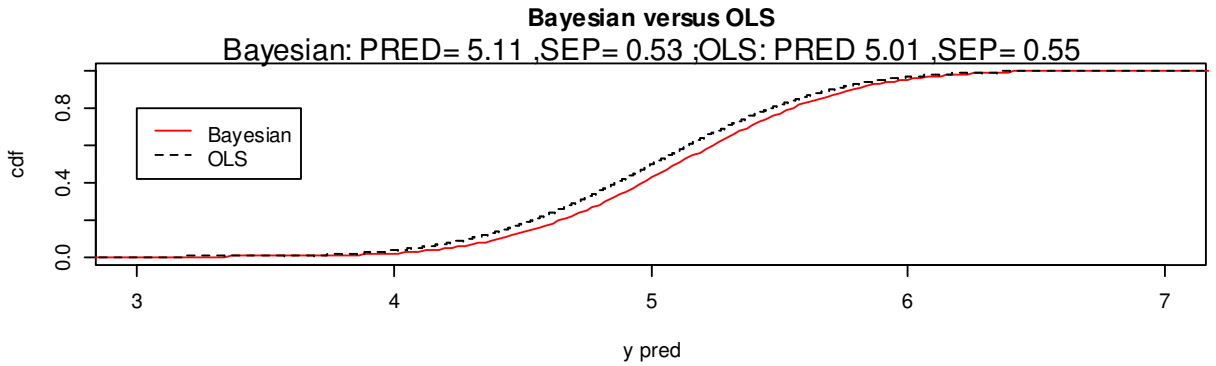


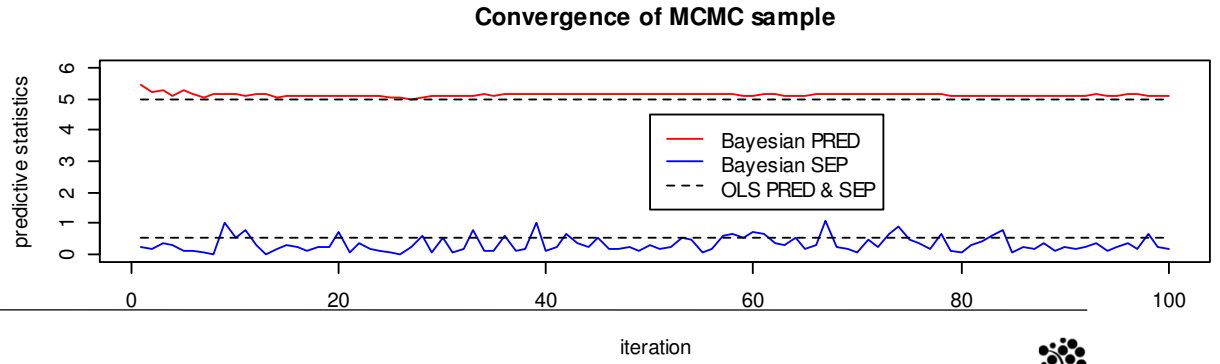


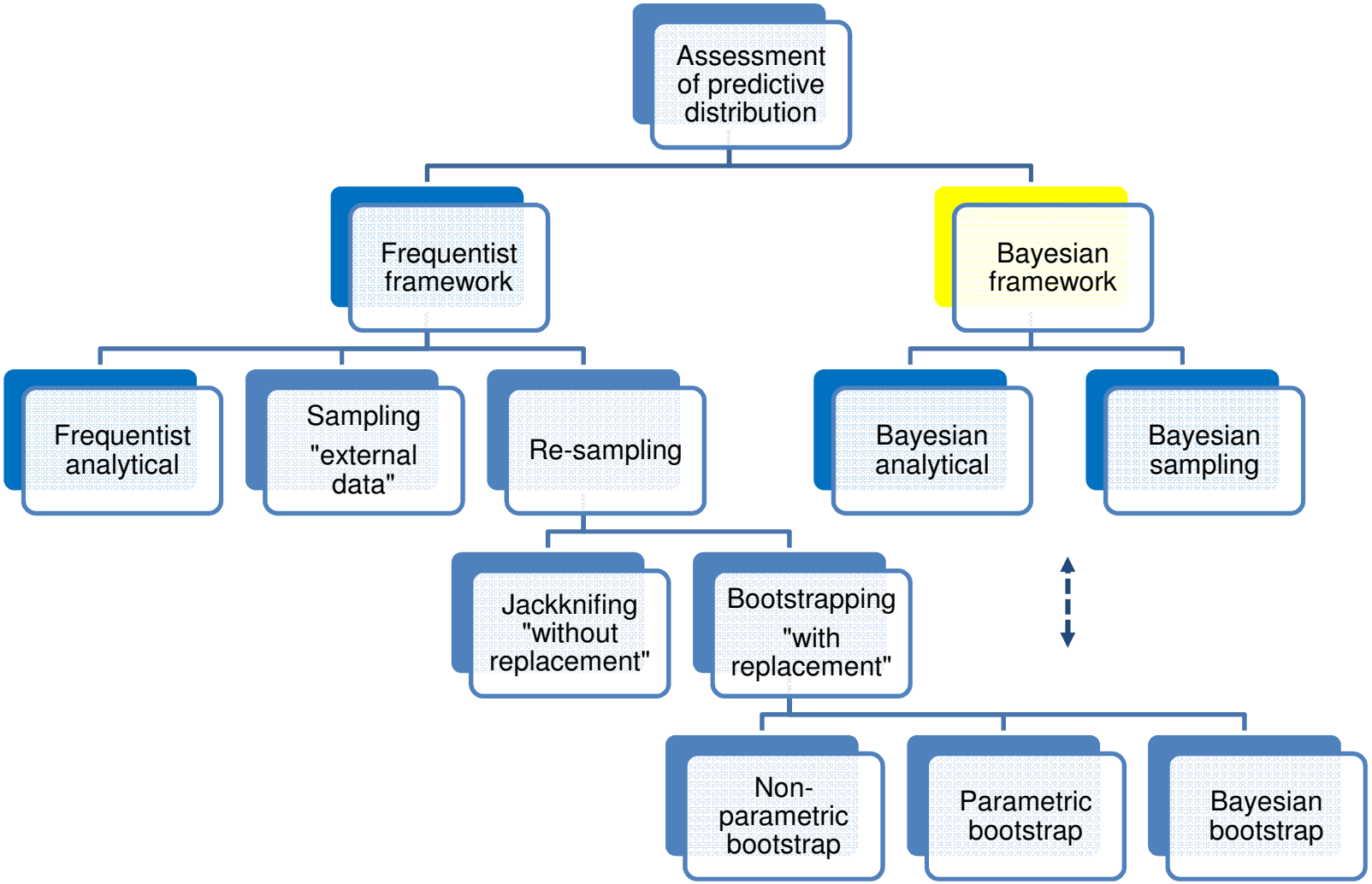
Sample from Bayesian with Normal priors

vs

Frequentist analytical



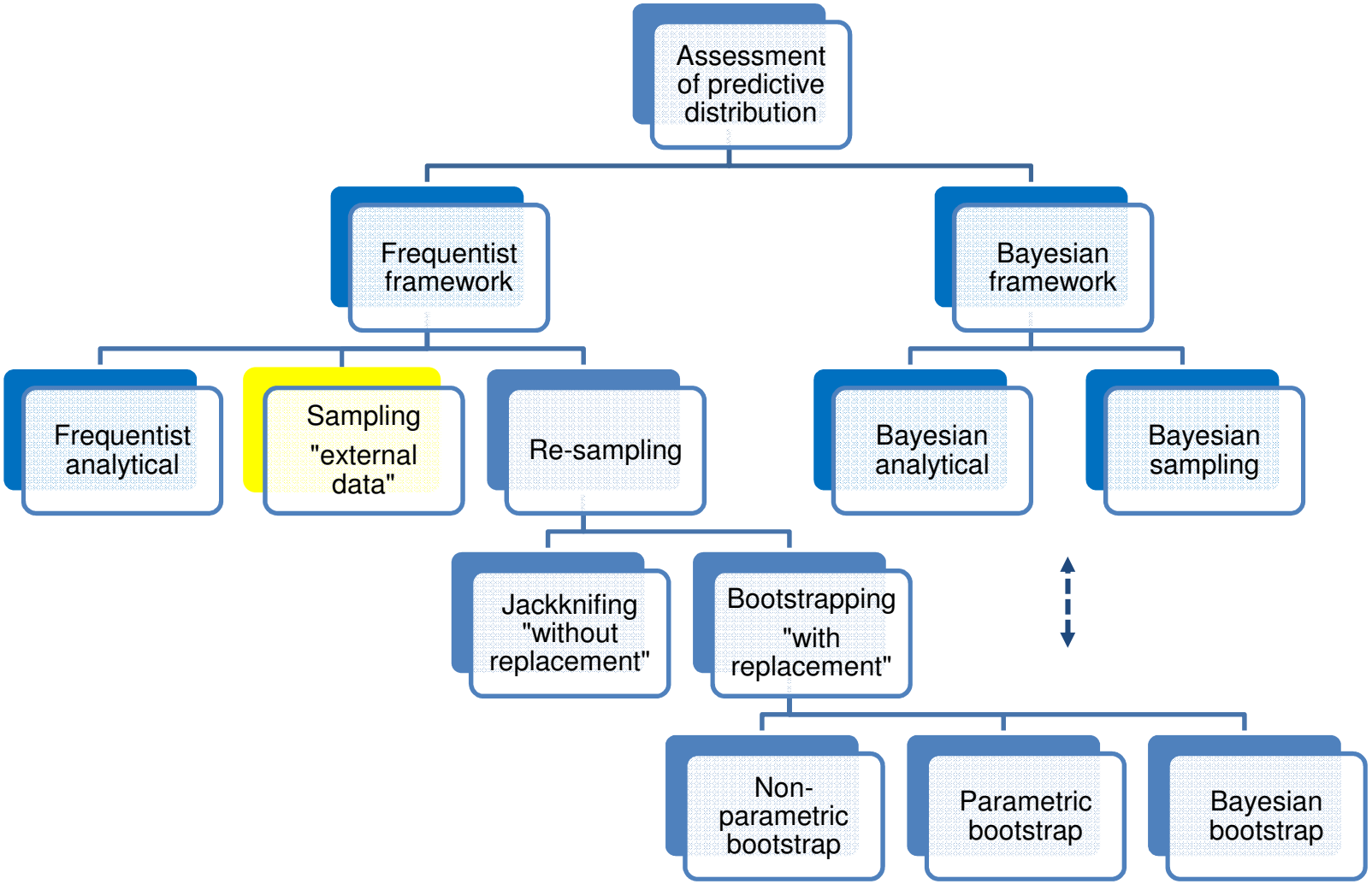



Assessment of predictive distribution
Frequentist framework
Bayesian framework
Frequentist analytical
Sampling "external data"
Re-sampling
Bayesian analytical
Bayesian sampling
Jackknifing "without replacement"
Bootstrapping "with replacement"
Non-parametric bootstrap
Parametric bootstrap
Bayesian bootstrap

Expert judgment

Based on experience of experimental data
Fingertop know how

Kind of distribution
Size of errors or coefficient of variation


Assessment of predictive distribution
Frequentist framework
Bayesian framework
Frequentist analytical
Sampling "external data"
Re-sampling
Bayesian analytical
Bayesian sampling
Jackknifing "without replacement"
Bootstrapping "with replacement"
Non-parametric bootstrap
Parametric bootstrap
Bayesian bootstrap

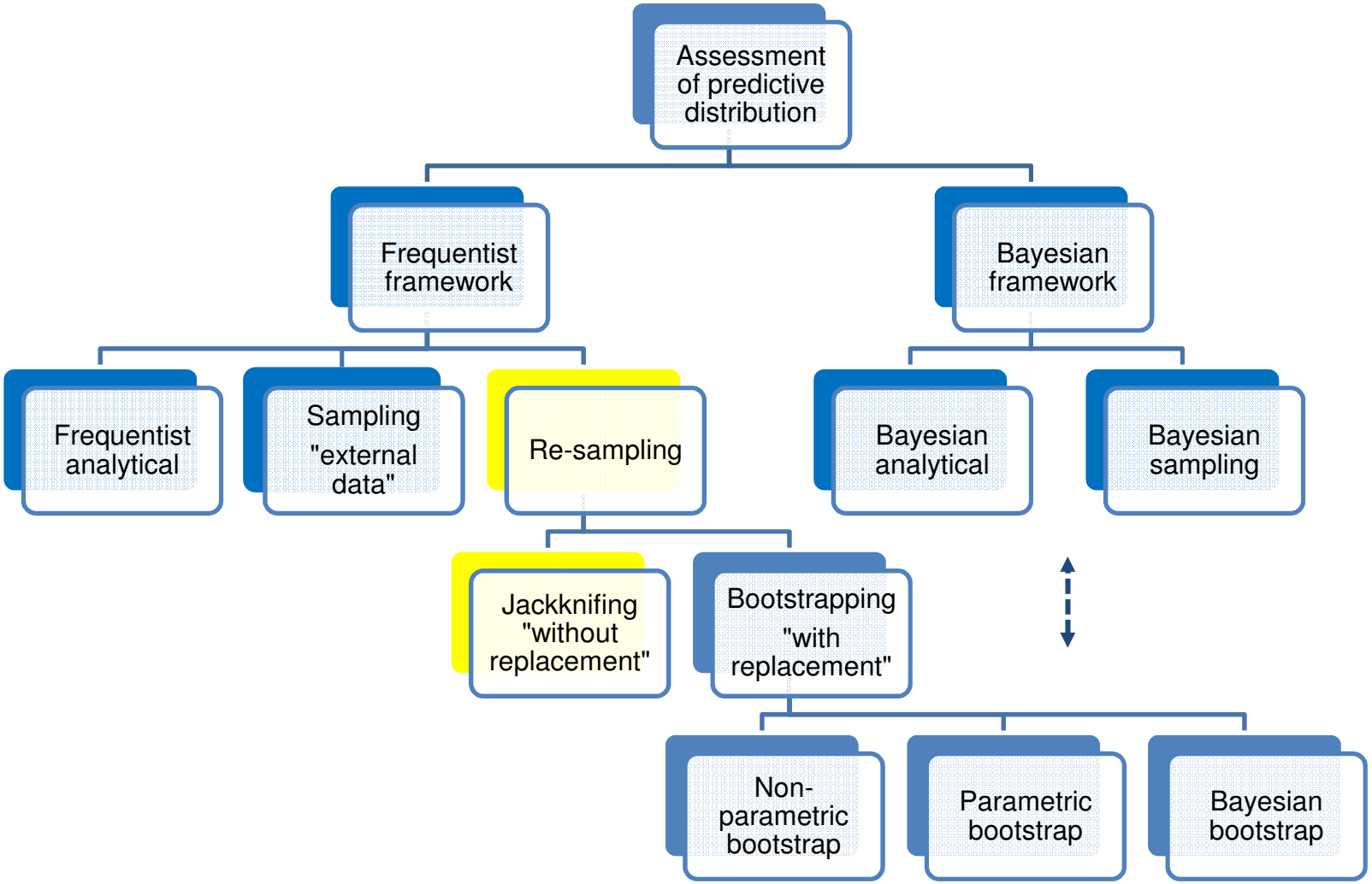
PRedictive Error Sum of Squares (PRESS)

Sampling: External data set

SIMPLE RULE OF THUMB : Gaussian + PRESS/nExt + point prediction

Is the predictive distribution really Gaussian?

Isnt it unlikely to have equal error and distribution over the AD?


Assessment of predictive distribution
Frequentist framework
Bayesian framework
Frequentist analytical
Sampling "external data"
Re-sampling
Bayesian analytical
Bayesian sampling
Jackknifing "without replacement"
Bootstrapping "with replacement"
Non-parametric bootstrap
Parametric bootstrap
Bayesian bootstrap

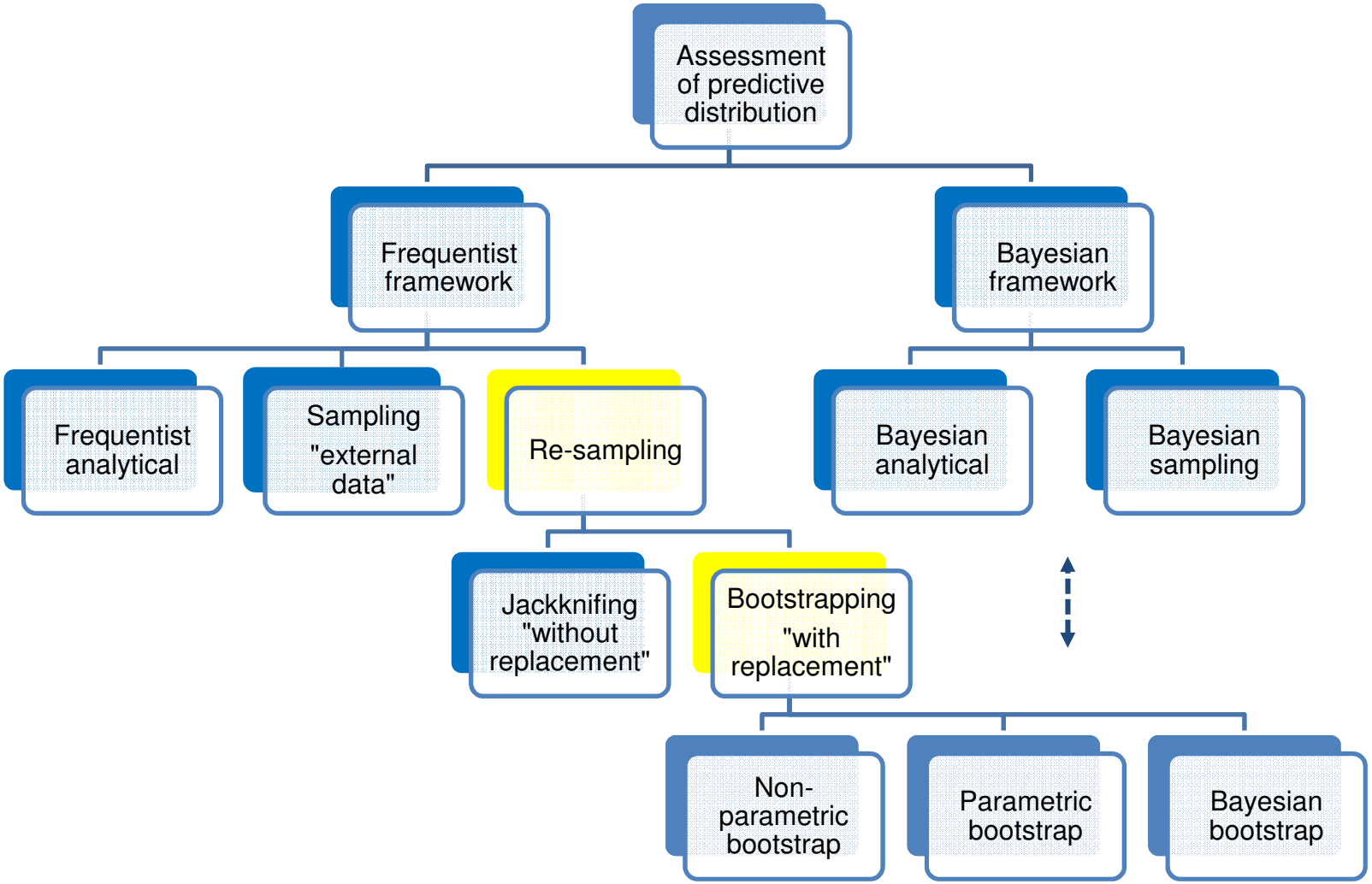
Predictive Error Sum of Squares (PRESS)

Re-sampling: Leave-One-Out PRESS

LOO PRESS divide by d.f. to get the variance of predictive error. What is the d.f. when we dont have a parametric model?

Is possible to asses without d.f. using bootstrapping

Note. The assessor need access to each individual residual to assess the variance of the predictive distribution since it is based on what assumption that is made on the kind of probablity distribution used.


Assessment of predictive distribution
Frequentist framework
Bayesian framework
Frequentist analytical
Sampling "external data"
Re-sampling
Bayesian analytical
Bayesian sampling
Jackknifing "without replacement"
Bootstrapping "with replacement"
Non-parametric bootstrap
Parametric bootstrap
Bayesian bootstrap

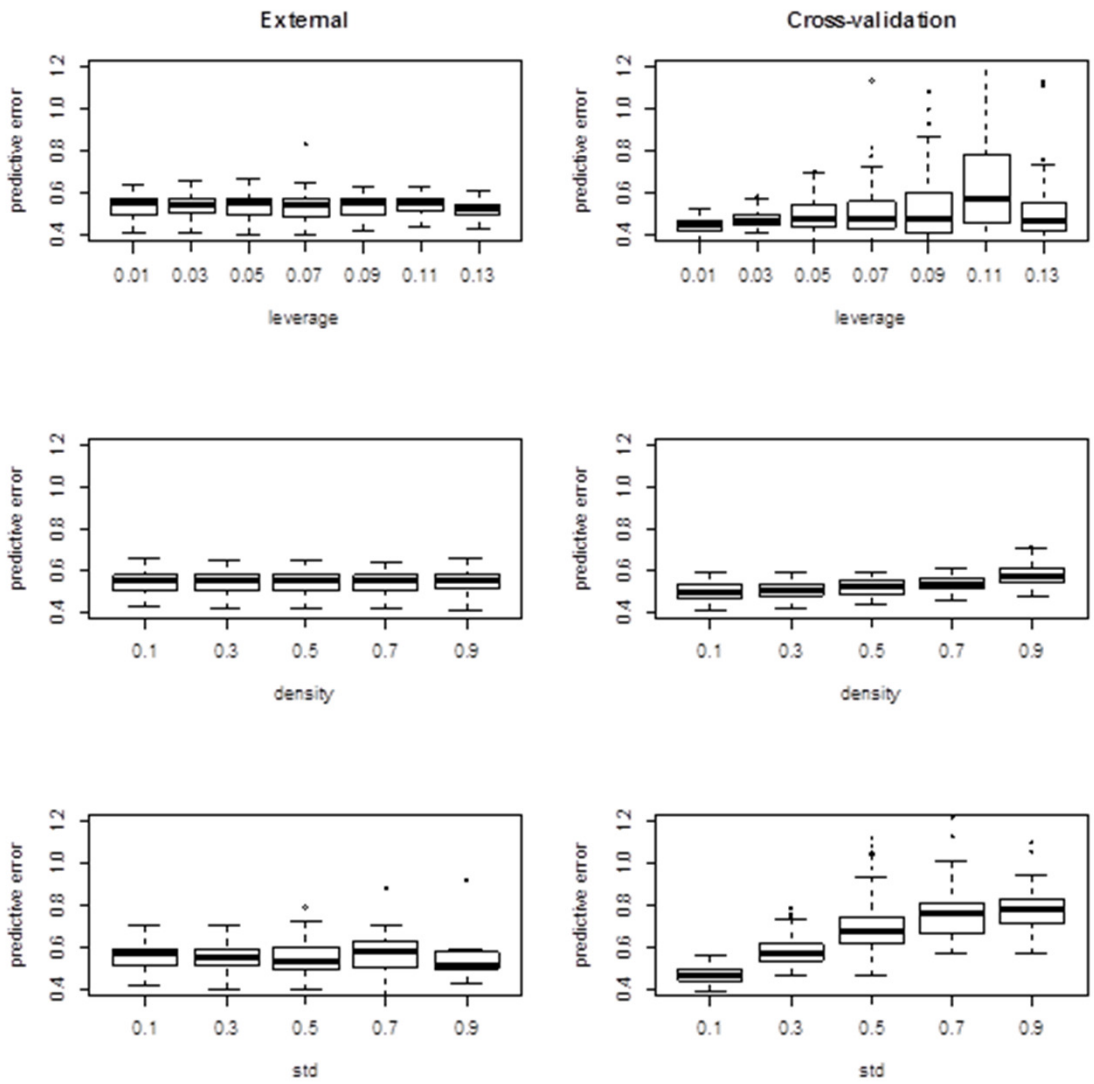
Localized PRESS

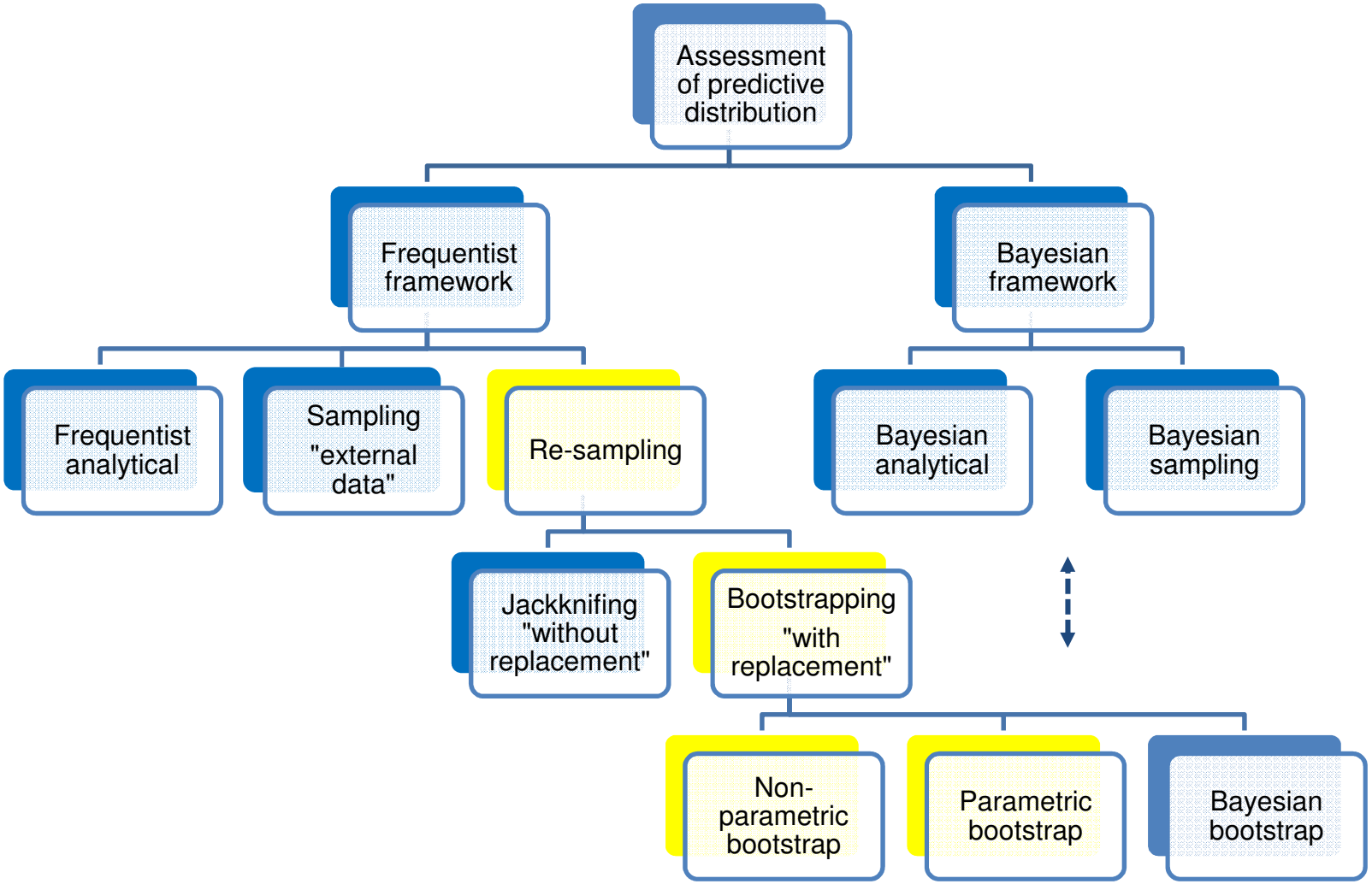
Variance = function of AD

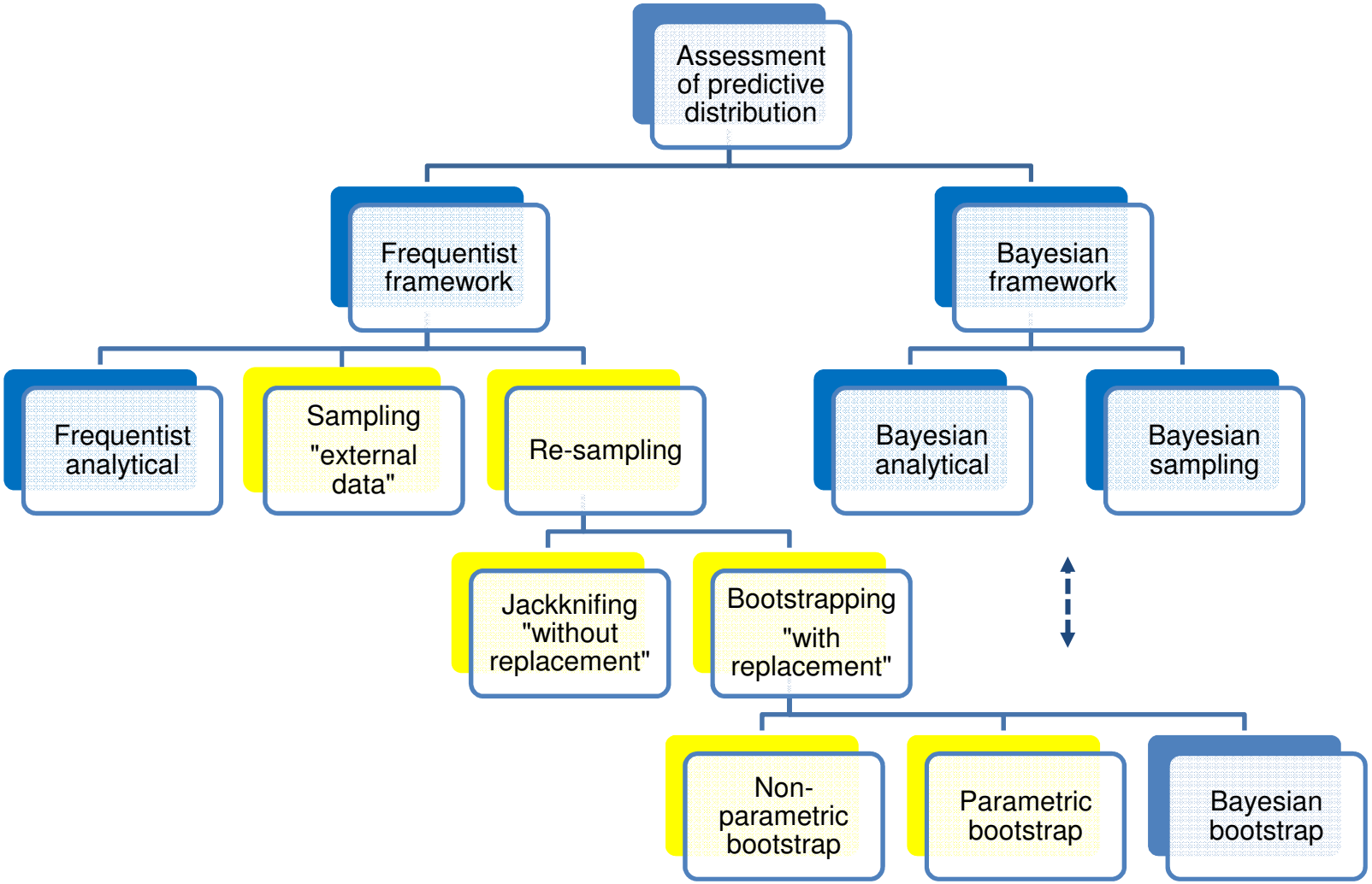
Modified residuals

Assume parametric distribution or not

Assessor need access each individual predictive residual or the complete data set and supervised learning algorithm

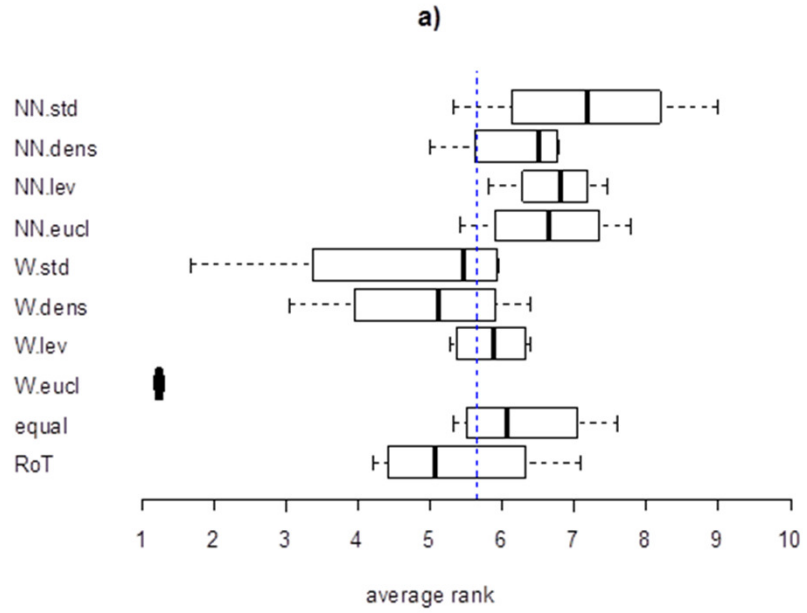


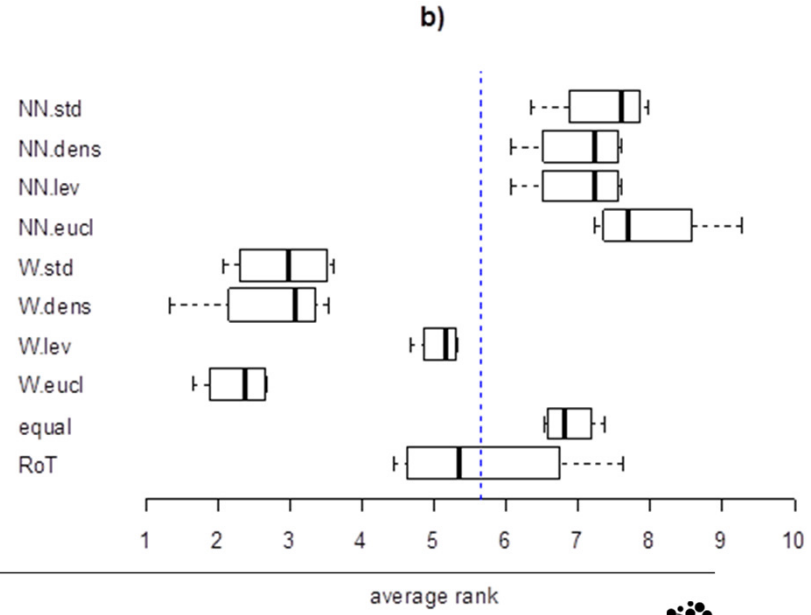

Assessment of predictive distribution
Frequentist framework
Bayesian framework
Frequentist analytical
Sampling "external data"
Re-sampling
Bayesian analytical
Bayesian sampling
Jackknifing "without replacement"
Bootstrapping "with replacement"
Non-parametric bootstrap
Parametric bootstrap
Bayesian bootstrap


Assessment of predictive distribution
Frequentist framework
Bayesian framework
Frequentist analytical
Sampling "external data"
Re-sampling
Bayesian analytical
Bayesian sampling
Jackknifing "without replacement"
Bootstrapping "with replacement"
Non-parametric bootstrap
Parametric bootstrap
Bayesian bootstrap

Evaluation

- NN-PRESS better than weighted PRESS and Rule of Thumb
- Non-parametric better than Gaussian
- Similarty by Euclidean distances and standard deviation in perturbed predictions top AD measures





Guidance

Wiki for approaches

Suggestions on how to evaluate which methods that are good

Understanding why and how we assess uncertainty

For exampel interpretation depends on perspective

Risk manager – Bayesian interpretation – this is used in a decision analysis, it is his/hers uncertainty given current background knowledge

Risk assessor – Balance between producing a precise but honest characterisation of uncertianty, is under pressure, needs to fullfill requirements for risk assessment, unc should be his/her uncertainty and thus have a Bayesian interpreation

QSAR modeller / **risk assessor** – ”this is an objective estimate of the uncertainty for this particular QSAR model” but ”I have chosen this method to assess it becuase …”

Communication: QSAR Model Reporting Format – given the need to assess predictive uncertainty

Continous update of QSARs

A validated QSAR serve as reference for waht performance measures that are good enough, then allow validated QSAR data to be entered to the model and check if the new performance measrues are good enough

Can a Bayesian model be accepted even though we cannot show predictive measures?

Important for small QSAR data

The algorithm to select descriptors should be part of the model

it is important to get honest predictive errors in re-sampling procedures

Communication: QSAR Prediction Reporting Format – given the need to assess predictive uncertainty

3.2.d. ”Predicted value (model result)” – point prediction view!

3.4. ”If possible, comment on the uncertainty of the prediction for this chemical, taking into account relevant information (e.g. variabilty of the experimental results)”.

Minimum requirements (such as report error in prediction in terms of a probability distribution).

Variability of experimental data is not considered when making the prediction

do we need new QSARs considering variability

It is the assessors uncertainty that are to be reflected.

Conclusions

Uncertainty is an associated characteristic of a QSAR prediction

Predictive uncertainty consist of predictive reliability and predictive error

For the purpose of uncertainty analysis using probabilities the predictive error is to be specified by a probability distribution

There are several approaches to assess predictive uncertainty

There will be subjectivity in the choice of approach to assess predictive uncertainty

Guidance

- Description of assessment approaches
- Methods to evaluate assessment uncertainty in QSAR prediction
- Communication of uncertainty in QSAR predictions

the end