**Types of problems for the final**

1.

1. For a given model, perform sensitivity analysis and remove random variables that do not affect the output much
2. Propagate uncertainties through the model
3. Replace the model with the surrogate model using Gaussian processes
4. Propagate uncertainties through the new model

EXAMPLE

Uncertainty propagation and sensitivity analysis (from OpenTurns)

<http://openturns.github.io/openturns/master/examples/uncertainty_propagation/estimate_probability_monte_carlo.html>

This model is a simple beam, restrained at one side and stressed by a traction load F at the other side.

The geometry is supposed to be deterministic: the diameter D is fixed to D=20 mm.

It is considered that failure occurs when the beam plastifies, i.e. when the axial stress gets bigger than the yield stress$\sigma\_e = \frac{F}{\pi-D^2/4} \leq 0$

Therefore, the state limit G used here is: $G = \sigma\_e - \frac{F}{\pi-D^2/4}$

Two independent random variables R and S are considered:

* R (the strength): $R = \sigma\_e$
* S (the load): $S = \frac{F}{\pi-D^2/4}$

Stochastic model:

* F ~ Normal(75e3, 5e3) [N]
* R ~ LogNormal(300, 30) [N]

1. Compute probability that G is less than 0.
2. Analyze sensitivity of the result to the parameters R and S.

2. You are developing a new device and for that device you performed measurements and compared it against the reference. Both reference and data and data from you device are given

1. Develop a model for that device – Bayesian regression
2. Now, the physical model is given – calibrate that model.
3. Data from another device on the market is given – draw scatter and Bland Altman plots.
4. Fit data to the mixed effect model
   1. Use traditional methods as we did in the lecture
   2. Use Bayesian based on PyMC3
   3. Find the agreement metrics and comment on them
5. Now, the device is in production and it is based on the regression model. Given these inputs, compute the estimates together with their uncertainties.

EXAMPLE

1. Download data from <http://staff.pubhealth.ku.dk/~tag/Teaching/share/data/Bodyfat.html> . The goal is to compute body fat using the parameters that are measured
   1. Develop a model for that data – Bayesian regression using all 13 parameters starting from Age.
   2. Perform sensitivity analysis and determine which of these 13 parameters are important.
   3. Make two models with reduced number of parameters.
   4. Calculate the WAIC or DIC and perform model averaging. Does it improve the results?
   5. Next, obtain the results using the formula: Percentage of Body Fat (i.e. 100\*B) = 495/D – 450. Find the agreement between the regression model and formula based model.

This is based on the paper: Jennifer A. Hoeting, David Madigan, Adrian E. Raftery and Chris T. Volinsky, Bayesian Model Averaging: A Tutorial, Statistical Science, Vol. 14, No. 4, 382–417, 1999.

3. You are supposed to design a system and need to select the components for the system. The desired accuracy of the system is 1% of the nominal value. Select the components of the system.

4. Data is given. Perform hierarchical/mixure modeling and compare performance of different models. Do model averaging in PyMC3. See the example in PyMC3: A Primer on Bayesian Methods for Multilevel Modeling <https://docs.pymc.io/notebooks/multilevel_modeling.html#Example:-Radon-contamination-(Gelman-and-Hill-2006)>

5. Time series data is given. Define AR(2) model. Analyze the uncertainty of the parameters of the model. Perform prediction using that model. See the example: Bayesian Auto-Regressive Time Series Analysis in PYMC3 <http://barnesanalytics.com/bayesian-auto-regressive-time-series-analysis-pymc3>

6. State-space model is given. It is nonlinear and non Gaussian. Analyze performance of that system using particle filters. How confident are you in your predictions.

7. Kalman filter

EXAMPLE 1: <https://arxiv.org/ftp/arxiv/papers/1204/1204.0375.pdf>

EXAMPLE 2: This problem is related to voltage measurement of the battery of the car. Every time the voltage is measured it has different value due to noise. Filtering of the noise will be done using Kalman filter. The model is as follows:

$x\_{k+1}=x\_k$

$z\_k=x\_k+v\_k$

X\_0=14, $v\_k is from Normal (0,2)$

* 1. Generate noisy data for N=50 measurements
  2. Perform Kalman filtering
  3. Analyze the error covariance as the time passes

8. Fit data to the distribution

Example – generate data from a normal distribution Normal(4.0, 1.5). Add noise. Estimate parameters of the distribution.