Chapter 5
Model Checking

Xavi Puig Oriol
Departament d'Estadística i I.O.

Chapter 5. Model Checking

Bayesian Model

$$\tilde{y} \mid \theta \sim p(\tilde{y} \mid \theta) \qquad \theta \sim \pi(\theta) \longrightarrow p(\tilde{y})$$

Data → **Bayes' Th.** / **MCMC**

$$\pi(\theta \mid y) = \frac{L_y(\theta)\pi(\theta)}{\int L_y(\theta)\pi(\theta)d\theta}$$

MCMC → Algorithms that simulate values from the posterior distribution

Bayesian Model (updated)

$$\tilde{y} \mid \theta \sim p(\tilde{y} \mid \theta) \qquad \theta \mid y \sim \pi(\theta \mid y) \longrightarrow p(\tilde{y} \mid y)$$

The updated model allows you to calculate point estimations, interval estimations, hypothesis test, predictions…

Chapter 5. Model Checking

The four steps of statistics

1. Design of the experiment
- Design how to collect the Data
- Choosing the statistical model
- Choosing the prior distribution

2. Model Checking

3. Statistical Inference
- Point Estimation
- Interval Estimation
- Prediction
- .....

4. Results presentation
- Showing and interpreting the results

Chapter 5. Model Checking

2. Model Checking

After choosing the model $M = \{p(\tilde{y} \mid \theta), \theta \in \Omega\}$

and collecting the data $y_1, y_2, \ldots, y_n$

You have to decide if the model is valid or not.

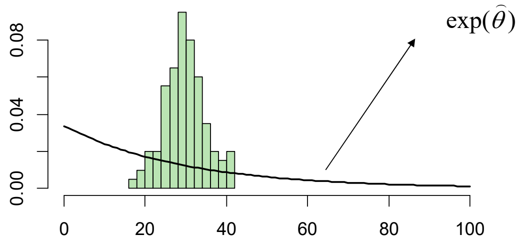
That is, you have to decide if the probability model that has generated the data belongs to M and therefore the model is correct.

Chapter 5. Model Checking

Example 1

Model Chosen $M = \{p(\tilde{y} \mid \theta) = \exp(\theta), \theta \in \Re_{+}\}$

Data Colected $y_1 = 20.2, y_2 = 35.7, \ldots, y_{100} = 27.4$



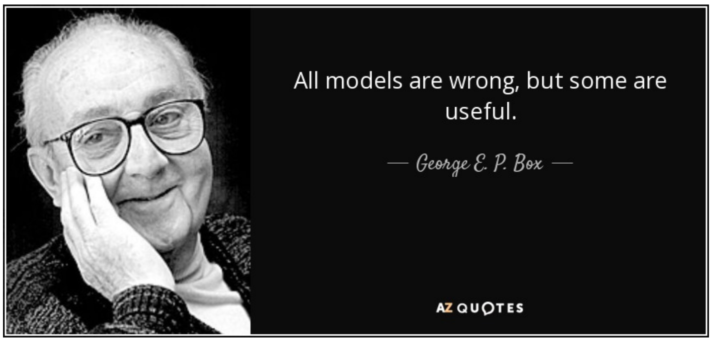
It is unlikely that Data has been generated by an Exponential Model. Therefore, the Model is not Valid.

Chapter 5. Model Checking

Main deas:

- The model will be valid if it is able to generate data like the observed data.
- The Bayesian model is a "data simulator", as it uses the posterior predictive distribution to simulate data from it.

Chapter 5. Model Checking



From Wikipedia:

George E. P. Box (1919 – 2013) was a British statistician, who worked in the areas of quality control, time-series analysis, design of experiments, and Bayesian inference. He has been called "one of the great statistical minds of the 20th century".

Chapter 5. Model Checking

Steps to check the model:

1. Choose a statistic(s)
2. Compute its reference distribution under the model
3. Compare the statistic with the reference distribution

Chapter 5. Model Checking

Steps to check the model:

1. Choose a statistic(s)

Numerical $T(y=(y_1,y_2,\ldots,y_n))$ or graphical

that summarizes the data and focuses on those relevant aspects to our objective.

The statistic(s) is ad-hoc to the problem. All the checking process is ad-hoc to the problem we are analyzing.

Chapter 5. Model Checking

Steps to check the model:

2. Compute the reference distribution of the statistic under the model

Using the posterior predictive distribution, we will simulate values of the statistic to approximate the reference distribution

we will simulate replicas of the observed data set

$$y^{rep}=(y_1^{rep},y_2^{rep},\ldots,y_n^{rep}) \quad \text{for} \quad rep=1,\ldots,M$$

and for each replica we will calculate the statistic,

$$T(y^{rep})$$

these values will allow us to approximate the reference distribution,

$$p(T(\tilde{y})\mid y)$$

Chapter 5. Model Checking

Steps to check the model:

3. Compare the statistic with the reference distribution

Numerically, for example by calculating the area of the tail,

$$\min\{p(T(\tilde{y}) \leq T(y) \mid y), p(T(\tilde{y}) \geq T(y) \mid y)\}$$

or graphically.

Evaluate the compatibility of the observed data with the model

Chapter 5. Model Checking

Example 2: sea dike

Chapter 5. Model Checking

Example 2: sea dike

For the design of a dike it is necessary to know the behavior of the waves, and one of the important factors is the height of the waves.

H is one of the parameters required for the design of a dike, where H means the probability of the wave of maximum height in a day exceeding the height H is less than 1e-06,

$$p(\tilde{y} > H \mid y) < 0.000001$$

The objective of the study is to estimate *H*.

The data available are the maximum daily wave heights observed during one year.

$\tilde{y}_i :=$ the maximum wave height on *i-th* day

Chapter 5. Model Checking

Example 2: sea dike

Proposed Model:

$$\tilde{y}_i \mid \mu \sim Normal(\mu, \sigma)$$
$$\mu \sim Normal(1, 4)$$
$$\sigma \sim Uniform(0, 10)$$

Observed data: $y_1 = 1.1, y_2 = 0.3, \ldots, y_{365} = 4.6$

```
> summary(y)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  0.100   0.200   0.500   1.015   1.400   7.500
```

Chapter 5. Model Checking

Example 2: sea dike

Step 1: Choose a statistic(s) that summarizes the data and is focused on those aspects that we consider relevant for our objective

1. $T_{\max}(y=(y_1,y_2,\ldots,y_n))=\max(y_1,y_2,\ldots,y_n)$

2. $T_{mean}(y=(y_1,y_2,\ldots,y_n))=mean(y_1,y_2,\ldots,y_n)$

Chapter 5. Model Checking

Example 2: sea dike

Step 2: Compute the reference distribution of the statistic under the model

replicas

simulations

$$
\begin{array}{cc|c|cc}
\mu & \sigma & & & \\
\mu^{(1)} & \sigma^{(1)} & \longrightarrow\ \tilde{y}_1^{(1)},\ldots,\tilde{y}_{365}^{(1)} \longrightarrow & T_{\max}(\tilde{y}^{(1)}) & T_{mean}(\tilde{y}^{(1)}) \\
\mu^{(2)} & \sigma^{(2)} & \longrightarrow\ \tilde{y}_1^{(2)},\ldots,\tilde{y}_{365}^{(2)} \longrightarrow & T_{\max}(\tilde{y}^{(2)}) & T_{mean}(\tilde{y}^{(2)}) \\
\vdots & & \vdots & \vdots & \\
\vdots & & \vdots & \vdots & \\
\mu^{(M)} & \sigma^{(M)} & \longrightarrow\ \tilde{y}_1^{(M)},\ldots,\tilde{y}_{365}^{(M)} \longrightarrow & T_{\max}(\tilde{y}^{(M)}) & T_{mean}(\tilde{y}^{(M)})
\end{array}
$$

$$\tilde{y}_i^{(rep)} \sim Norm(\mu^{(rep)},\sigma^{(rep)})$$

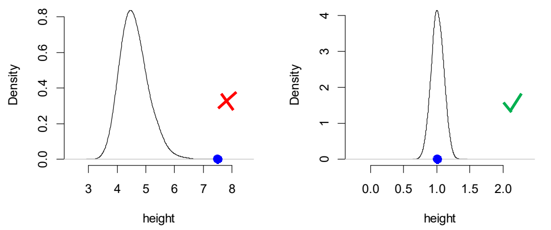
Chapter 5. Model Checking

Example 2: sea dike

Step 3: Compare the observed statistic with the reference distribution ($T_{\max}(y) = 7.5, \quad T_{mean}(y) = 1.015$)

Graphically

$p(T_{\max}(\tilde{y}) \mid y)$ $p(T_{mean}(\tilde{y}) \mid y)$



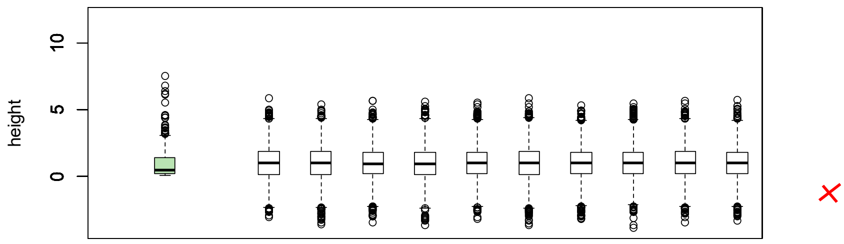
Numerically

$$\min\{p(T_{\max}(\tilde{y}) \le 7.5 \mid y), p(T_{\max}(\tilde{y}) \ge 7.5 \mid y)\} < 0.000$$
$$\min\{p(T_{mean}(\tilde{y}) \le 1.015 \mid y), p(T_{mean}(\tilde{y}) \ge 1.015 \mid y)\} = 0.492$$

Chapter 5. Model Checking

Example 2: sea dike

Other graphical validation

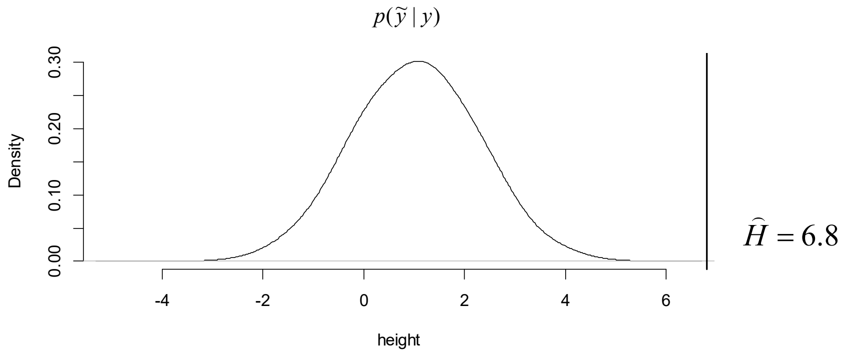


Chapter 5. Model Checking

Example 2: sea dike

H estimation using this not valid model

$$p(\tilde{y} > H \mid y) < 0.000001$$



In the observed sample of 365 days there are 2 days in which the waves exceeded 6.8 meters, when by definition the threshold of H should only be exceeded once per million days!!

Chapter 5. Model Checking

Example 2: sea dike

New proposed model:

$$\tilde{y}_i \mid \alpha, \beta \sim Gamma(\alpha, \beta) \qquad \text{for i =1,…,365}$$
$$\alpha \sim Gamma(0.01, 0.01)$$
$$\beta \sim Gamma(0.01, 0.01)$$

Step 1: Choose a statistic(s) that summarizes the data and is focused on those aspects that we consider relevant for our objective

1. $T_{\max}(y = (y_1, y_2, \ldots, y_n)) = \max(y_1, y_2, \ldots, y_n)$
2. $T_{mean}(y = (y_1, y_2, \ldots, y_n)) = mean(y_1, y_2, \ldots, y_n)$

Chapter 5. Model Checking

Example 2: sea dike

Step 2: Compute the reference distribution of the statistic under the model

$$
\begin{array}{c|cc|c|c|c|cc|}
 & \alpha & \beta & & \text{replicas} & & & \\
\text{simulations} & \alpha^{(1)} & \beta^{(1)} & \longrightarrow & \tilde{y}_1^{(1)}, \ldots, \tilde{y}_{365}^{(1)} & \longrightarrow & T_{\max}(\tilde{y}^{(1)}) & T_{mean}(\tilde{y}^{(1)}) \\
 & \alpha^{(2)} & \beta^{(2)} & \longrightarrow & \tilde{y}_1^{(2)}, \ldots, \tilde{y}_{365}^{(2)} & \longrightarrow & T_{\max}(\tilde{y}^{(2)}) & T_{mean}(\tilde{y}^{(2)}) \\
 & \vdots & & & \vdots & & \vdots & \\
 & \vdots & & & \vdots & & \vdots & \\
 & \alpha^{(M)} & \beta^{(M)} & \longrightarrow & \tilde{y}_1^{(M)}, \ldots, \tilde{y}_{365}^{(M)} & \longrightarrow & T_{\max}(\tilde{y}^{(M)}) & T_{mean}(\tilde{y}^{(M)})
\end{array}
$$

$$\tilde{y}_i^{(rep)} \sim Gamma(\alpha^{(rep)}, \beta^{(rep)})$$

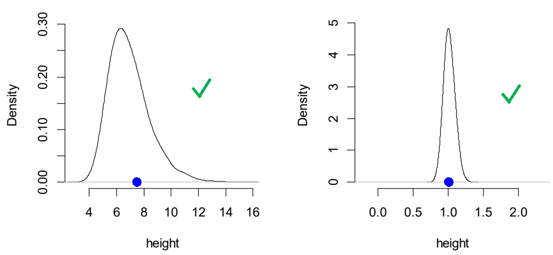
Chapter 5. Model Checking

Example 2: sea dike

Step 3: Compare the observed statistic with the reference distribution ($T_{\max}(y) = 7.5, \quad T_{mean}(y) = 1.015$)

Graphically

$p(T_{\max}(\tilde{y}) \mid y)$ $\qquad$ $p(T_{mean}(\tilde{y}) \mid y)$



Numerically

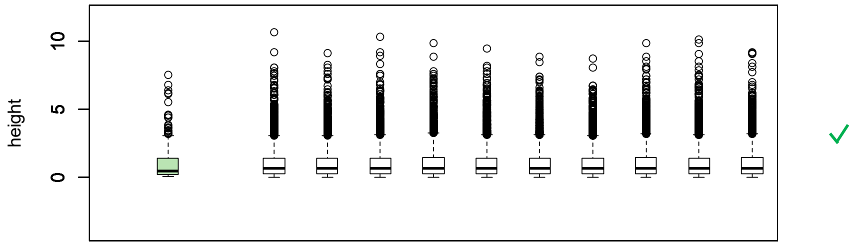
$$\min\{p(T_{\max}(\tilde{y}) \le 7.5 \mid y), p(T_{\max}(\tilde{y}) \ge 7.5 \mid y)\} = 0.299$$ ✓

$$\min\{p(T_{mean}(\tilde{y}) \le 1.015 \mid y), p(T_{mean}(\tilde{y}) \ge 1.015 \mid y)\} = 0.495$$ ✓

Chapter 5. Model Checking

Example 2: sea dike

Other graphical validation



Observed Data

10 replicas

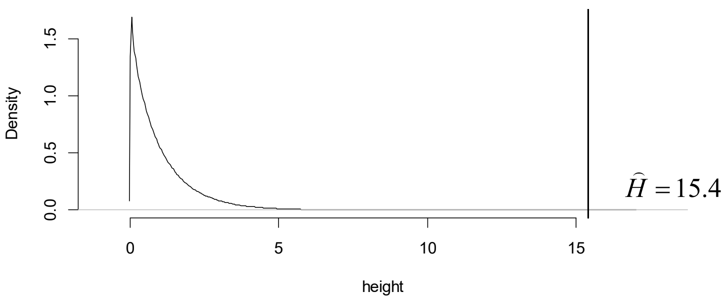
Chapter 5. Model Checking

Example 2: sea dike

H estimation using this valid model

$$p(\tilde{y} > H \mid y) < 0.000001$$

$p(\tilde{y} \mid y)$



In the observed sample of 365 days, no day has exceeded the threshold of H=15.4, the maximum observed was 7.5m.

Chapter 5. Model Checking

If the objective is to make a prediction, a common way to validate the model is:

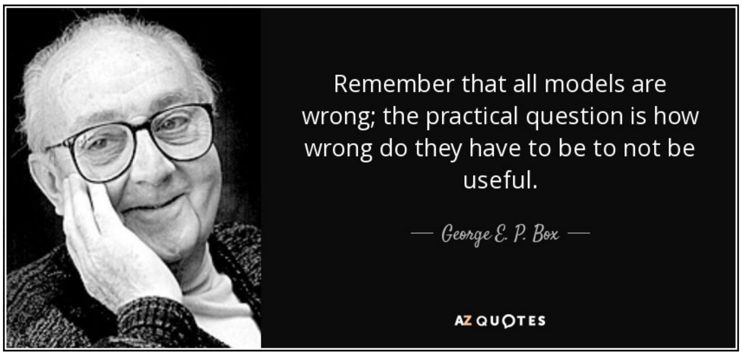
1. Extract a subset of the data
2. Implement the model without the extracted data
3. Compare the predictions the model makes for the extracted data with this extracted data (for example, noting whether the prediction intervals contain the observed data).

Chapter 5. Model Checking

Model Construction

The iterative process of proposing a model, validating it, identifying its limitations and perhaps suggesting a new model is what we call Model Construction.

Chapter 5. Model Checking



Chapter 5. Model Checking

I encourage you to read chapter 6 (*Model Checking and Improvement*) from the book:

Gelman A, Carlin J, Stern H, Dunson D, Vehtari A, and Rubin D (2014). Bayesian Data Analysis (3rd ed). London: Chapman & Hall.

Chapter 5
Model Checking

Xavi Puig Oriol
Departament d'Estadística i I.O.