Analyses in Bayesian Networks

Uffe Kjærulff (uk@cs.aau.dk)

Group of Machine Intelligence
Department of Computer Science, Aalborg University

Reykjavik University, 29 April, 2005

Outline

1. Data conflict analysis
2. Value of information analysis
3. Sensitivity analysis relative to observations
4. Sensitivity analysis relative to parameters

Data Conflict Analysis

- What is data conflict?
- Data conflict measure
- Tracing conflicts
- Conflict or rare case?

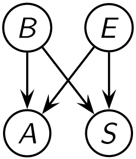
Data Conflict

- Inconsistencies among observations are easily detected ($P(\varepsilon) = 0$).
- Negatively correlated observations can lead to opposing hypotheses, neutralizing each others effect on a hypothesis variable.
- A flawed observation will be negatively correlated with non-flawed observations.
- Flawed observation should be detected and traced.
- In a diagnostic situation a single flawed test result may take the investigation in a completely wrong direction.
- Rare case: A Bayesian network represents a closed world with a finite set of variables and causal relations (holds true only under certain assumptions).
- Alert the user if a set of observations is not well covered by the model.

Seismometer

Dr. Watson makes frequent calls to Mr. Holmes regarding the *burglar* alarm. Every time Mr. Holmes rushes home, just to find that everything is in order, since, till now, the cause of activation of the *alarm* has been small *earthquakes*. So now Mr. Holmes is installing a *seismometer* in his house with a direct line to his office.

S: No, Small, and Large vibrations.



Mr. Holmes

One afternoon Dr. Watson calls again and announces that Mr. Holmes' alarm has gone off. Mr. Holmes checks the seismometer, it is in state 0 (i.e., no vibrations).

- From our knowledge of the model, we would say that the findings are in conflict.
- A propagation does not disclose the conflict ($P(B = \text{yes}) = 0.38$).

Using the model only, we cannot distinguish between flawed data and a case not covered by the model.

The Conflict Measure

We need a conflict measure that is easy to calculate and gives an indication of a possible conflict. Two pieces of evidence $\varepsilon = \{\varepsilon_i, \varepsilon_j\}$ are

- positively correlated if $P(\varepsilon_i | \varepsilon_j) > P(\varepsilon_i)$,
- negatively correlated if $P(\varepsilon_i | \varepsilon_j) < P(\varepsilon_i)$,
- independent if $P(\varepsilon_i | \varepsilon_j) = P(\varepsilon_i)$.

The Conflict Measure

We need a conflict measure that is easy to calculate and gives an indication of a possible conflict. Two pieces of evidence $\varepsilon = \{\varepsilon_i, \varepsilon_j\}$ are

- positively correlated if $P(\varepsilon_i | \varepsilon_j) > P(\varepsilon_i)$,
- negatively correlated if $P(\varepsilon_i | \varepsilon_j) < P(\varepsilon_i)$,
- independent if $P(\varepsilon_i | \varepsilon_j) = P(\varepsilon_i)$.

There is an indication of a conflict between ε_i and ε_j, if

$$\frac{P(\varepsilon_i)P(\varepsilon_j)}{P(\varepsilon_i, \varepsilon_j)} > 1 \Longleftrightarrow \log \frac{P(\varepsilon_i)P(\varepsilon_j)}{P(\varepsilon_i, \varepsilon_j)} > 0$$

The Conflict Measure

We need a conflict measure that is easy to calculate and gives an indication of a possible conflict. Two pieces of evidence $\varepsilon = \{\varepsilon_i, \varepsilon_j\}$ are

- positively correlated if $P(\varepsilon_i | \varepsilon_j) > P(\varepsilon_i)$,
- negatively correlated if $P(\varepsilon_i | \varepsilon_j) < P(\varepsilon_i)$,
- independent if $P(\varepsilon_i | \varepsilon_j) = P(\varepsilon_i)$.

There is an indication of a conflict between ε_i and ε_j, if

$$\frac{P(\varepsilon_i)P(\varepsilon_j)}{P(\varepsilon_i, \varepsilon_j)} > 1 \iff \log \frac{P(\varepsilon_i)P(\varepsilon_j)}{P(\varepsilon_i, \varepsilon_j)} > 0$$

The conflict between ε_i and ε_j is

$$\mathrm{conf}(\varepsilon) = \log \frac{P(\varepsilon_i)P(\varepsilon_j)}{P(\varepsilon)}.$$

The Conflict Measure

- Let $\varepsilon = \{\varepsilon_1, \ldots, \varepsilon_n\}$ be a set of observations (evidence).
- For positively correlated findings we expect that

$$P(\varepsilon) > \prod_{i=1}^{n} P(\varepsilon_i)$$

Thus, the conflict measure is defined as

$$\mathrm{conf}(\varepsilon) = \mathrm{conf}(\{\varepsilon_1, \ldots, \varepsilon_n\}) = \log \frac{\prod_{i=1}^{n} P(\varepsilon_i)}{P(\varepsilon)}.$$

A positive $\mathrm{conf}(\varepsilon)$ indicates a possible conflict.

Mr. Holmes

The evidence $\varepsilon = \{S = 0, A = \text{yes}\}$.

$$
\begin{aligned}
\text{conf}(\varepsilon) &= \text{conf}(\{S = 0, A = \text{yes}\}) \\
&= \log \frac{P(S = 0)P(A = \text{yes})}{P(S = 0, A = \text{yes})} \\
&= \log \frac{0.44 \cdot 0.55}{0.012} \\
&= 3.0 \\
&> 0.
\end{aligned}
$$

Thus, $\text{conf}(\varepsilon)$ indicates a possible conflict.

Conflict or Rare Case?

Typical data from a very rare case may indicate a possible conflict.

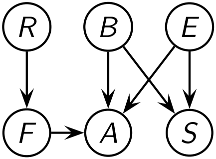
Let $\varepsilon = \{\varepsilon_1, \ldots, \varepsilon_n\}$ be findings for which $\mathrm{conf}(\varepsilon) > 0$ and let h be a hypothesis which could explain the findings (i.e. $\mathrm{conf}(\varepsilon \cup \{h\}) \leq 0$):

$$\begin{aligned}\mathrm{conf}(\varepsilon \cup \{h\}) &= \log \frac{P(\varepsilon_1) \cdots P(\varepsilon_n) P(h)}{P(\varepsilon, h)} \\ &= \mathrm{conf}(\varepsilon) + \log \frac{P(h)}{P(h \,|\, \varepsilon)}.\end{aligned}$$

Thus, if $\mathrm{conf}(\varepsilon) \leq \log \dfrac{P(h \,|\, \varepsilon)}{P(h)}$, then h can explain away the conflict (*normalized likelihood*).

Holmes looks out his window. It rains cats and dogs.

- F: Flood, R: Rain, $\varepsilon = \{R = \text{heavy}, S = 0, A = \text{yes}\}$.



The posterior probability of flood is $P(F = \text{yes} \mid \varepsilon) = 0.99$ and prior is $P(F = \text{yes}) = 0.006$ — the conflict is explained away as a rare case.

- The conflict is $\text{conf}(\varepsilon) = -0.24$.

Tracing Conflicts

After the conflict measure has been found to indicate a possible conflict, the conflict should be traced.

- Compute the conflict measure for different subsets ε' of ε.

In the example, we have three subsets with partial conflicts:

- $\text{conf}(\{\varepsilon_A, \varepsilon_R\}) = -0.45$
- $\text{conf}(\{\varepsilon_A, \varepsilon_S\}) = 3.03$
- $\text{conf}(\{\varepsilon_R, \varepsilon_S\}) = 0$
- Local conflict: $\text{conf}(\{\{\varepsilon_A, \varepsilon_R\}, \varepsilon_S\}) = 0.213$.

Global conflict is the sum of the local and partial conflicts:

$$\text{conf}(\{\varepsilon_A, \varepsilon_R, \varepsilon_S\}) = \text{conf}(\{\varepsilon_A, \varepsilon_R\}) + \text{conf}(\{\{\varepsilon_A, \varepsilon_R\}, \varepsilon_S\})$$

Summary

- Data conflict
- Data conflict measure
- Tracing conflicts
- Conflict or rare case

Outline

1. Data conflict analysis
2. Value of information analysis
3. Sensitivity analysis relative to observations
4. Sensitivity analysis relative to parameters

Value of Information Analysis in Bayesian Networks

- Value of information analysis
- Myopic value of information analysis
- Value of information analysis in influence diagrams
- Non-myopic value of information analysis
- Value of information analysis in a Bayesian network

Value of Information

Before deciding on an action more information can be acquired.

- Seldomly cost free.
- Is it worthwhile consulting additional information sources.
- If more than one source exists, the task is to come up with a strategy for consulting the information sources.

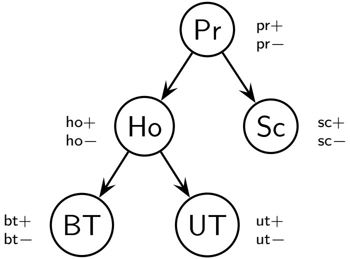
Additional information (if free) cannot make you worse off.

No value of information if you will not change your decision.

Insemination

Insemination

Six weeks after insemination of a cow there are three tests for the result: *blood test* (BT), *urine test* (UT), and *scanning* (Sc). The results of the blood test and the urine test are mediated through the *hormonal state* (Ho) which is affected by a possible *pregnancy* (Pr).



Insemination

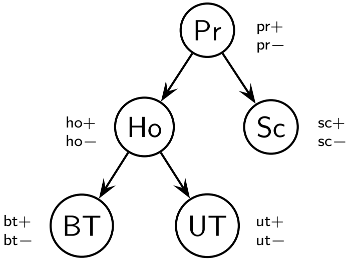
- Assume that you have the options to *repeat* the insemination or to *wait* for another six-weeks period.
- The cost of repeating the insemination is 65 units no matter the pregnancy state of the cow. If the cow is pregnant, and you wait, it will cost you nothing, but if the cow is not pregnant, and you wait, it will cost you an additional 30 units plus the eventual repeated insemination (that makes a total of 95 units for waiting).

	wait	repeat
pr+	0	−65
pr−	−95	−65

- A *blood test* has a cost of 1 unit and a *urine test* has a cost of 2 units.

Value of Information

Hypothesis driven data request.



- The value of the information scenario with respect to hypothesis Pr is:

$$V_{\mathrm{Pr}} = \max_{a \in A} \sum_{h \in \mathrm{Pr}} U(a,h)P(h).$$

- A proper analysis of the data request situation consists of an analysis of all possible sequences.

Myopic Value of Information

Assume we are allowed to consult at most one information source.

- If test T with cost C_T yields outcome t, then the value of the new information scenario is

$$V_{\mathrm{Pr}}(t) = \max_{a \in A} \sum_{h \in \mathrm{Pr}} U(a, h) P(h \mid t).$$

- Since the outcome of T is not known we calculate the expected value

$$\mathrm{EV}_{\mathrm{Pr}}(T) = \sum_{t \in T} V_{\mathrm{Pr}}(t) P(t).$$

- The expected benefit is $\mathrm{EB}_{\mathrm{Pr}}(T) = \mathrm{EV}_{\mathrm{Pr}}(T) - V_{\mathrm{Pr}}$.
- The expect profit is $\mathrm{EP}_{\mathrm{Pr}}(T) = \mathrm{EB}_{\mathrm{Pr}}(T) - C_T$.

Insemination

With $P(\text{pr+}) = 0.87$, $P(\text{pr+} \,|\, \text{bt+}) = 0.976$, $P(\text{pr+} \,|\, \text{bt}-) = 0.729$, and $P(\text{bt+}) = 0.571$ we get

$$\begin{aligned}
V_{\text{Pr}} &= \max_{a \in \{\text{wait},\text{repeat}\}} \sum_{h \in \{\text{pr+},\text{pr}-\}} U(a,h)P(h) \\
&= \max\{U(\text{wait},\text{pr+})P(\text{pr+}) + U(\text{wait},\text{pr}-)P(\text{pr}-), \\
&\qquad U(\text{repeat},\text{pr+})P(\text{pr+}) + U(\text{repeat},\text{pr}-)P(\text{pr}-)\} \\
&= \max\{0 \cdot 0.87 + (-95) \cdot 0.13, -65 \cdot 0.87 + (-65) \cdot 0.13\} \\
&= -12.35,
\end{aligned}$$

which is the value associated with the hypothesis variable Pr with nothing observed.

Insemination

Now, if BT is observed, we get

$$\begin{aligned}
V_{\mathrm{Pr}}(\mathrm{bt}+) &= \max_{a\in\{\mathrm{wait},\mathrm{repeat}\}} \sum_{h\in\{\mathrm{pr}+,\mathrm{pr}-\}} U(a,h)P(h \,|\, \mathrm{bt}+) \\
&= \max\{0 \cdot 0.976 + (-95)\cdot 0.0.024, \\
&\qquad -65 \cdot 0.976 + (-65)\cdot 0.024\} \\
&= -2.28,
\end{aligned}$$

$$\begin{aligned}
V_{\mathrm{Pr}}(\mathrm{bt}-) &= \max\{0 \cdot 0.729 + (-95)\cdot 0.0.271, \\
&\qquad -65 \cdot 0.729 + (-65)\cdot 0.271\} \\
&= -25.75.
\end{aligned}$$

Insemination

Then if we weigh these values with the probabilities of the associated observations, we get

$$\begin{aligned}\mathrm{EV}_{\mathsf{Pr}}(BT) &= \sum_{t\in\{\mathrm{bt}+,\mathrm{bt}-\}} V_{\mathsf{Pr}}(t)P(t)\\ &= -2.28 \cdot 0.571 + (-25.75) \cdot 0.429 = -12.35,\end{aligned}$$

which is exactly the same as $\mathrm{V}_{\mathsf{Pr}}(BT)$! Hence

$$\mathrm{EB}_{\mathsf{Pr}}(BT) = \mathrm{EV}_{\mathsf{Pr}}(BT) - V_{\mathsf{Pr}} = 0,$$

and

$$\mathrm{EP}_{\mathsf{Pr}}(BT) = \mathrm{EB}_{\mathsf{Pr}}(BT) - C_{\mathsf{BT}} = -1.$$

Insemination: Interpretation of Result

We found

$$V_{\mathsf{Pr}} = \mathrm{EV}_{\mathsf{Pr}}(BT) = -12.35,$$

and hence

$$\mathrm{EB}_{\mathsf{Pr}}(BT) = \mathrm{EV}_{\mathsf{Pr}}(BT) - V_{\mathsf{Pr}} = 0,$$

meaning that we do not gain anything by getting the extra information from a blood test.

> *Value of information*
>
> The value of information is zero, unless it will make you change your decision.

Insemination: Interpretation of Result

In particular, we found

$$\arg \max_{a \in \{\text{wait}, \text{repeat}\}} \sum_{h \in \{\text{pr}+, \text{pr}-\}} U(a, h) P(h \mid \text{bt}+) = \text{wait}$$

$$\arg \max_{a \in \{\text{wait}, \text{repeat}\}} \sum_{h \in \{\text{pr}+, \text{pr}-\}} U(a, h) P(h \mid \text{bt}-) = \text{wait}$$

That is, no matter the outcome of BT our decision should be "wait".

An interesting question:

> How much can the model parameters (i.e., the probabilities and utilities) change without affecting the result of the analysis?

Answer can be provided through *sensitivity analysis*.

Non-Myopic Value of Information

- Assume we are allowed to consult any number of information sources.
- Important if the expected benefit of consulting a pair of information sources is greater than the sum of the costs.
- This is a much more computationally involved task to perform.
- If costs cannot be reduced by performing tests simultaneously, then deciding to perform two tests can never be better than to consult one information source and decide afterwards whether to consult the second.

VOI Analysis in Bayesian Networks

- How do we perform value of information analysis without specifying utilities?
- The reason for acquiring more information is to decrease the uncertainty about a hypothesis.
- The *entropy* is a measure of how much probability mass is scattered around on the states (the degree of chaos).

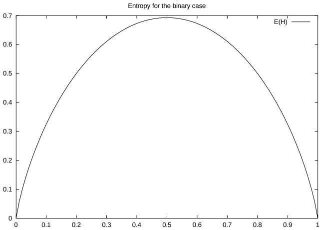
$$H(P(H)) = -\sum_{h \in H} P(h) \log_2(P(h)).$$

- Thus, $H(P(H)) \in [0, \log_2(n)]$ where $|H| = n$.
- Entropy is a measure of randomness. The more random a variable is, the higher its entropy.

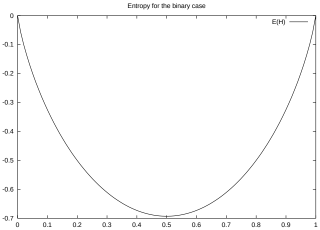
Entropy

- If the entropy is to be used as a value function, then

$$V_H = -H(P(H)) = \sum_{h \in H} P(h) \log_2(P(h)).$$



$H(P(H))$



$V_H = -H(P(H))$

- We want to maximize $V_H = -H(P(H))$ (i.e., minimize $H(P(H))$).

Entropy

- What is the expected most informative observation?
  - The conditional entropy is

$$H(T \mid X) = -\sum_X P(X) \sum_T P(T \mid X) \log_2 P(T \mid X).$$

- Let T be the target, now select X with maximum information gain

$$MI(T, X) = H(T) - H(T \mid X) = H(X) + H(T) - H(X, T)$$

- A measure of the reduction of the entropy of T given X.

Summary

- Value of information analysis
- Myopic value of information analysis
- Value of information analysis in influence diagrams
- Non-myopic value of information analysis
- Value of information analysis in Bayesian networks

Outline

1. Data conflict analysis
2. Value of information analysis
3. Sensitivity analysis relative to observations
4. Sensitivity analysis relative to parameters

Sensitivity Analysis

Given a model N and a hypothesis variable H we would like to determine the sensitivity of the model or hypothesis relative to the observations made or the parameters of the model.

- Sensitivity analysis with respect to ε can give answers to questions like:
  - Which evidence is in favor of/against/irrelevant for h_i.
  - Which evidence discriminate h_i from h_j?

- A structural analysis can give answers to some of these questions, but this is not the point here.

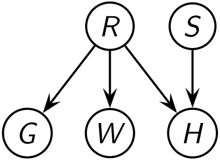
Wet Grass

In the morning when Mr. Holmes leaves his house he realizes that his grass is wet. He wonders whether it has rained during the night or whether he has forgotten to turn off his sprinkler. He looks at the grass of his neighbors, Dr. Watson and Mrs. Gibbon. Both lawns are dry and he concludes that he must have forgotten to turn off his sprinkler.

$$h_S : S = \text{yes and } \varepsilon = \{\varepsilon_G, \varepsilon_W, \varepsilon_H\}$$

Mr. Holmes: Structure

The structure:



Which pieces of evidence are Mr. Holmes' reasoning sensitive to?

- Recall *d*-separation.

Mr. Holmes: Sensitivity Analysis

- We have $P(h_S) = 0.1$ and $P(h_S \mid \varepsilon) = 0.9999$.
- Since $P(h_S \mid \varepsilon_H) = 0.51$, $P(h_S \mid \varepsilon_W) = 0.1 = P(h_S \mid \varepsilon_G)$, we conclude:
  - Neither ε_W nor ε_G alone have any impact on h_S.
  - ε_H is not sufficient for the conclusion.
- The conclusion that ε_W and ε_G are irrelevant is not correct.
- The evidence in combination has a larger impact than the "sum" of the individual impacts.

Sensitivity Analysis Concepts

Some loosely defined concepts:

- Let $\varepsilon = \{\varepsilon_1, \ldots, \varepsilon_n\}$ be a set of observations, and let $\varepsilon' \subseteq \varepsilon$.
- ε' is *sufficient* if $P(h|\varepsilon')$ is *almost equal* to $P(h|\varepsilon)$.
  - $\varepsilon \setminus \varepsilon'$ is *redundant*.
- ε' is *minimally* sufficient, if it is sufficient and no $\varepsilon'' \subset \varepsilon'$ is.
- ε' is *crucial*, if it is a subset of all sufficient sets.
- ε' is *important* if $P(h|\varepsilon)$ change too much without it.

Now, let's try to be more precise.

Sufficiency and Importance

- Let $\varepsilon = \{\varepsilon_1, \ldots, \varepsilon_n\}$ be a set of observations, and let $\varepsilon' \subseteq \varepsilon$.

- ε' is *sufficient* if $P(h \mid \varepsilon')$ is *almost equal* to $P(h \mid \varepsilon)$:
$$\left| \frac{p(h \mid \varepsilon')}{p(h \mid \varepsilon)} - 1 \right| < \theta_1.$$

- ε' is *important* if the probability of h change too much without it:
$$\left| \frac{p(h \mid \varepsilon \setminus \varepsilon')}{p(h \mid \varepsilon)} - 1 \right| > \theta_2.$$

Redundancy and Irrelevance

- We distinguish between redundancy and irrelevance.

- A subset $\varepsilon' \subseteq \varepsilon$ is *redundant* if $\varepsilon \setminus \varepsilon'$ is sufficient; i.e., if

$$\left| \frac{P(h \mid \varepsilon \setminus \varepsilon')}{P(h \mid \varepsilon)} - 1 \right| < \theta.$$

- If two subsets of evidence ε' and ε'' are redundant, then both cannot necessarily be removed (e.g. wet grass example).

- A piece of evidence x is *irrelevant* for h if it is redundant in all subsets of ε:

$$\left| \frac{P(h \mid \varepsilon' \setminus \{x\})}{P(h \mid \varepsilon')} - 1 \right| < \theta, \ \forall \varepsilon' \subseteq \varepsilon.$$

Mr. Holmes: Redundancy and Irrelevance

Consider ε_G and ε_W in the Mr. Holmes example.

- $P(h_S \,|\, \varepsilon = \{\varepsilon_W, \varepsilon_G, \varepsilon_H\}) = 0.9999$
- $P(h_S \,|\, \varepsilon_G, \varepsilon_H) = P(h_S \,|\, \varepsilon_W, \varepsilon_H) = 0.988$
- $P(h_S \,|\, \varepsilon_H) = 0.51$

With $\theta = 0.02$ both ε_G and ε_W are redundant, but none of them are irrelevant.

What-if

- Let $\varepsilon = \{\varepsilon_1, \ldots, \varepsilon_n\}$ be a set of observations and assume a single hypothesis h is of interest.
- What if the observation ε_i had not been made, but ε'_i instead ?
- Involves computing $P(h \,|\, \varepsilon \cup \{\varepsilon'_i\} \setminus \{\varepsilon_i\})$ and comparing results.
- This kind of analysis will help you determine, if a subset of evidence acts for or against a hypothesis.

Discrimination of Hypotheses

- Let $H = (h_1, \ldots, h_m)$ be the hypotheses of interest.
- Question: Which evidence discriminate h_i from h_j?
- To relate the impact of ε' on h_i and h_j we can use:

$$\frac{P(\varepsilon' \,|\, h_i)}{P(\varepsilon' \,|\, h_j)} \quad = \quad \frac{L(h_i \,|\, \varepsilon')}{L(h_j \,|\, \varepsilon')}.$$

Discrimination of Hypotheses

Mr. Holmes has two hypotheses h_R and h_S and evidence $\varepsilon = \{\varepsilon_W, \varepsilon_G, \varepsilon_H\}$.

ε'			$\frac{P(\varepsilon' \mid h_S)}{P(\varepsilon' \mid h_R)}$
ε_G	ε_W	ε_H	6622
ε_G	ε_W	—	7300
ε_G	—	ε_H	74
ε_G	—	—	81
—	ε_W	ε_H	74
—	ε_W	—	81
—	—	ε_H	0.92
—	—	—	1

Thus, ε_G and ε_W are good discriminators.

Complexity of Sensitivity Analysis

- The heart of sensitivity analysis is the computation of

$$P(h \mid \varepsilon'),\ \forall \varepsilon' \subseteq \varepsilon,\ \forall h \in H.$$

- The complexity of this task grows exponentially.

Summary

- Classification of evidence as
  - sufficient
  - important
  - crucial
  - redundant
  - irrelevant
- What-if analysis.
- Discrimination among hypotheses.
- Sensitivity analyses relative to evidence can have high computational complexity.

Outline

Sensitivity Analysis Relative to Parameters

- Let N be a Bayesian network with parameters $\vec{t}$, where each $t \in \vec{t}$ is of the form $t = P(A = a \,|\, \mathrm{pa}(A) = \pi)$.
- A single hypothesis $H = h$ is of interest.
- We are interested in how $P(h \,|\, \varepsilon)$ varies with $\vec{t}$.
- It turns out that $P(\varepsilon)(t) = \alpha t + \beta$, $\alpha, \beta \in \mathbb{R}$. Thus

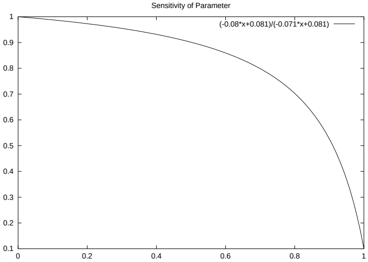
$$P(h \,|\, \varepsilon)(t) = \frac{P(h, \varepsilon)(t)}{P(\varepsilon)(t)} = \frac{\gamma t + \delta}{\alpha t + \beta}.$$

- The posterior probability is a fraction of two multi-linear functions of the parameters.
- The coefficients can easily be found through inference.

Determining The Sensitivity Function

In the Wet Grass example:

$$P(h_S \,|\, \varepsilon = \{\varepsilon_H, \varepsilon_G\})(t) = \frac{-0.08t + 0.081}{-0.071t + 0.081}$$



Whether or not the precision of an assessment of the value t_0 of a parameter t is important depends on the size of $|P'(h \,|\, \varepsilon)(t_0)|$.

The Derivative and Sensitivity Value

- The derivative of the sensitivity function tells us how much $P(h|\varepsilon)(t)$ change as a function of t.

- In general, $P'(h|\varepsilon)(t) = \dfrac{P(h|\varepsilon)(t)}{\partial t} = \dfrac{\alpha\delta - \beta\gamma}{(\gamma t + \delta)^2}$.

- The sensitivity value of a parameter t is $|P'(h|\varepsilon)(t_0)|$.
  - Where t_0 is the original assessment.
  - If $|P'(h|\varepsilon)(t_0)| > 0$, then t is of interest.

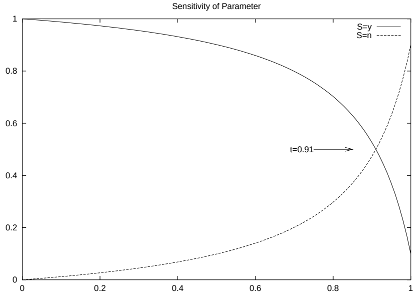
- In the example we for $t_0 = 0.1$ get $|P'(h_S|\varepsilon)(t_0)| = 0.08$.

- Not enough to consider the sensitivity value alone since approximation is good for small deviations only.

The Sensitivity Function

The alternative hypothesis $h_{S=n}$ has:

$$P(h_{S=n} \mid \varepsilon)(t) = \frac{0.009t}{-0.071t + 0.081}$$



For $t = 0.91$: $P(h_{S=n} \mid \varepsilon)(t) = P(h_{S=y} \mid \varepsilon)(t)$.

Which Parameters Should Be Investigated?

Parameters that deserve further investigation have:

- A sensitivity value $> \theta$.

How much can the parameter vary before the most likely hypothesis change (admissible deviation)?

- Sensitivity bounds $\Delta t \in [a, b]$ before hypothesis change.
- In the example, $\Delta t \in [-0.1, 0.91 - 0.1]$.

Summary

- The sensitivity function $P(h|\varepsilon)(t)$ for a hypothesis $H = h$ as a function of a parameter, say, $t = P(A = a|B = b)$ can easily be determined.
- A sensitivity function is a fraction of two (multi-)linear functions of the parameter(s).
- The derivative of a sensitivity function tells you how much $P(h|\varepsilon)(t)$ changes as a function of t.
- Sensitivity functions provide very useful information in the probability elicitation process, telling you with which precision you need to assess parameter values.
- *One-way* sensitivity analysis — how does $P(h|\varepsilon)$ change as a function of each parameter — is simple. *n-way* is computationally expensive.