Does Bayes Theorem Work?

Michael Goldstein
Durham University *

*Thanks for support from Basic Technology initiative (MUCM), NERC (RAPID), Leverhulme (Tipping Points)

RAPID-WATCH

What are the implications of RAPID-WATCH observing system data and other recent observations for estimates of the risk due to rapid change in the MOC? In this context risk is taken to mean the probability of rapid change in the MOC and the consequent impact on climate (affecting temperatures, precipitation, sea level, for example). This project must:

RAPID-WATCH

What are the implications of RAPID-WATCH observing system data and other recent observations for estimates of the risk due to rapid change in the MOC? In this context risk is taken to mean the probability of rapid change in the MOC and the consequent impact on climate (affecting temperatures, precipitation, sea level, for example). This project must:

* contribute to the MOC observing system assessment in 2011;
* investigate how observations of the MOC can be used to constrain estimates of the probability of rapid MOC change, including magnitude and rate of change;
* make sound statistical inferences about the real climate system from model simulations and observations;
* investigate the dependence of model uncertainty on such factors as changes of resolution;
* assess model uncertainty in climate impacts and characterise impacts that have received less attention (eg frequency of extremes).

The project must also demonstrate close partnership with the Hadley Centre.

Uncertainty in climate projections (from Met Office web-site)

1.1.1 What do we mean by probability in UKCP09?

Uncertainty in climate projections (from Met Office web-site)

1.1.1 What do we mean by probability in UKCP09?

It is important to point out early in this report that a probability given in UKCP09 (or indeed IPCC) is not the same as the probability of a given number arising in a game of chance, such as rolling a dice. It can be seen as the relative degree to which each possible climate outcome is supported by the evidence available, taking into account our current understanding of climate science and observations, as generated by the UKCP09 methodology. If the evidence changes in future, so will the probabilities.

Uncertainty in climate projections (from Met Office web-site)

1.1.1 What do we mean by probability in UKCP09?

It is important to point out early in this report that a probability given in UKCP09 (or indeed IPCC) is not the same as the probability of a given number arising in a game of chance, such as rolling a dice. It can be seen as the relative degree to which each possible climate outcome is supported by the evidence available, taking into account our current understanding of climate science and observations, as generated by the UKCP09 methodology. If the evidence changes in future, so will the probabilities.

Subjective probability is a measure of the degree to which a particular outcome is consistent with the information considered in the analysis (i.e. strength of the evidence) ... Probabilistic climate projections are based on subjective probability, as the probabilities are a measure of the degree to which a particular level of future climate change is consistent with the evidence considered. In the case of UKCP09, a Bayesian statistical framework was used, and the evidence comes from historical climate observations, expert judgement and results of considering the outputs from a number of climate models, all with their associated uncertainties.

Cosmic uncertainty

Cosmic uncertainty

Galaxy formation: a Bayesian Uncertainty Analysis Ian Vernon, Michael Goldstein and Richard G. Bower Bayesian Analysis (2010) 5, 619 - 67

ABSTRACT ... An uncertainty analysis of a computer model known as Galform is presented. Galform models the creation and evolution of approximately one million galaxies from the beginning of the Universe until the current day, and is regarded as a state-of-the-art model within the cosmology community. It requires the specification of many input parameters in order to run the simulation, takes significant time to run, and provides various outputs that can be compared with real world data.

Cosmic uncertainty

Galaxy formation: a Bayesian Uncertainty Analysis Ian Vernon, Michael Goldstein and Richard G. Bower Bayesian Analysis (2010) 5, 619 - 67

ABSTRACT ... An uncertainty analysis of a computer model known as Galform is presented. Galform models the creation and evolution of approximately one million galaxies from the beginning of the Universe until the current day, and is regarded as a state-of-the-art model within the cosmology community. It requires the specification of many input parameters in order to run the simulation, takes significant time to run, and provides various outputs that can be compared with real world data.

A Bayes Linear approach is presented in order to identify the subset of the input space that could give rise to acceptable matches between model output and measured data. This approach takes account of the major sources of uncertainty in a consistent and unified manner, including input parameter uncertainty, function uncertainty, observational error, forcing function uncertainty and structural uncertainty ...

Cosmic uncertainty

Galaxy formation: a Bayesian Uncertainty Analysis Ian Vernon, Michael Goldstein and Richard G. Bower Bayesian Analysis (2010) 5, 619 - 67

ABSTRACT ... An uncertainty analysis of a computer model known as Galform is presented. Galform models the creation and evolution of approximately one million galaxies from the beginning of the Universe until the current day, and is regarded as a state-of-the-art model within the cosmology community. It requires the specification of many input parameters in order to run the simulation, takes significant time to run, and provides various outputs that can be compared with real world data.

A Bayes Linear approach is presented in order to identify the subset of the input space that could give rise to acceptable matches between model output and measured data. This approach takes account of the major sources of uncertainty in a consistent and unified manner, including input parameter uncertainty, function uncertainty, observational error, forcing function uncertainty and structural uncertainty ...

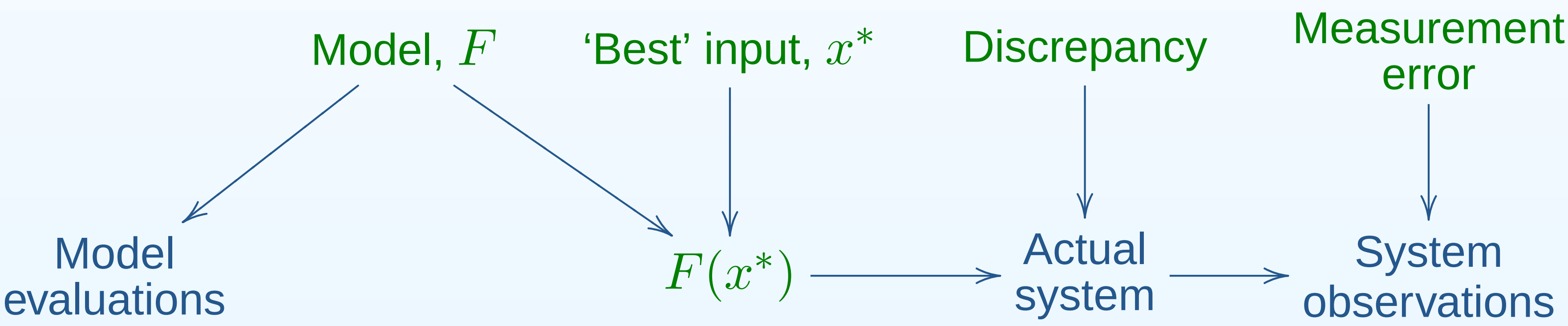
The analysis was successful in producing a large collection of model evaluations that exhibit good fits to the observed data.

Using models to quantify uncertainty

Modeller's fallacy Analysing the model is the same as analysing the system.

Using models to quantify uncertainty

Modeller's fallacy Analysing the model is the same as analysing the system. The most common way to 'correct' this fallacy is based on the idea that the model, F, is informative for system behaviour at the "best" input choice.



Using models to quantify uncertainty

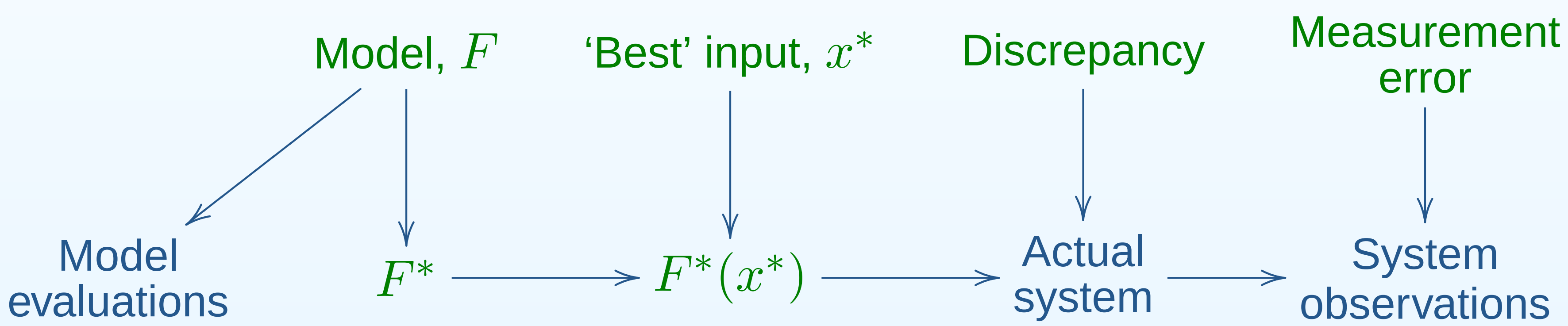
Modeller's fallacy Analysing the model is the same as analysing the system.

The most common way to 'correct' this fallacy is based on the idea that the model, F, is informative for system behaviour at the "best" input choice.

$$
\begin{tikzcd}
 & \text{Model, } F \arrow[dl] \arrow[d] & \text{'Best' input, } x^* \arrow[d] & \text{Discrepancy} \arrow[d] & \text{Measurement error} \arrow[d] \\
\text{Model evaluations} & F^* \arrow[r] & F^*(x^*) \arrow[r] & \text{Actual system} \arrow[r] & \text{System observations}
\end{tikzcd}
$$

Using models to quantify uncertainty

Modeller's fallacy Analysing the model is the same as analysing the system.
The most common way to 'correct' this fallacy is based on the idea that the model, F, is informative for system behaviour at the "best" input choice.



A model describes how system **properties** influence system **behaviour** simplifying both the properties and how they influence behaviour.
A full uncertainty representation must consider how model evaluations are informative for the actual relationship, F^*, [the "reified" model] between system properties and behaviour.
Now F^* is informative for system behaviour at the "best" input.

Some Questions

What do we mean by **uncertainty quantification**?

Some Questions

What do we mean by **uncertainty quantification**?

Does Bayes theorem quantify uncertainty?

[And if so, how?]

Some Questions

What do we mean by **uncertainty quantification**?

Does Bayes theorem quantify uncertainty?
[And if so, how?]

Alternately, is Bayes analysis actually a model for quantifying uncertainty?
[And if so, is it a good model?]

Some Questions

What do we mean by **uncertainty quantification**?

Does Bayes theorem quantify uncertainty?
[And if so, how?]

Alternately, is Bayes analysis actually a model for quantifying uncertainty?
[And if so, is it a good model?]

If Bayesian analysis is a model for uncertainty quantification, then do we need to correct for the modeller's fallacy, to bridge the gap between Bayesian model uncertainty quantification and real world quantification of uncertainty?
[And how could we do that?]

Subjectivist Bayes

In the subjectivist Bayes view, the meaning of any probability statement is straightforward. It is the uncertainty judgement of a specified individual, expressed on the scale of probability by consideration of some operational elicitation scheme, for example by consideration of betting preferences.

Subjectivist Bayes

In the subjectivist Bayes view, the meaning of any probability statement is straightforward. It is the uncertainty judgement of a specified individual, expressed on the scale of probability by consideration of some operational elicitation scheme, for example by consideration of betting preferences.

In the subjectivist interpretation, any probability statement is the judgement of a named individual, so we should speak not of the probability of rapid climate change, but instead of Anne's probability or Bob's probability of rapid climate change and so forth.

Subjectivist Bayes

In the subjectivist Bayes view, the meaning of any probability statement is straightforward. It is the uncertainty judgement of a specified individual, expressed on the scale of probability by consideration of some operational elicitation scheme, for example by consideration of betting preferences.

In the subjectivist interpretation, any probability statement is the judgement of a named individual, so we should speak not of the probability of rapid climate change, but instead of Anne's probability or Bob's probability of rapid climate change and so forth.

Most people expect something more authoritative and objective than a probability which is one person's judgement. However, the disappointing thing is that, in almost all cases, stated probabilities emerging from a complex analysis are not even the judgements of any individual.

Subjectivist Bayes

In the subjectivist Bayes view, the meaning of any probability statement is straightforward. It is the uncertainty judgement of a specified individual, expressed on the scale of probability by consideration of some operational elicitation scheme, for example by consideration of betting preferences.

In the subjectivist interpretation, any probability statement is the judgement of a named individual, so we should speak not of the probability of rapid climate change, but instead of Anne's probability or Bob's probability of rapid climate change and so forth.

Most people expect something more authoritative and objective than a probability which is one person's judgement. However, the disappointing thing is that, in almost all cases, stated probabilities emerging from a complex analysis are not even the judgements of any individual.

So, it is not unreasonable that an objective of our analysis should be probabilities which are asserted by at least one person (more would be good!).

Is this a sufficient objective?

Best current judgements

When is the probability of an individual scientifically valuable?

Best current judgements

When is the probability of an individual scientifically valuable?

[1] This individual is knowledgeable in the area

Best current judgements

When is the probability of an individual scientifically valuable?

[1] This individual is knowledgeable in the area

[2] the analysis that has led to this judgement has been sufficiently careful, thorough and exhaustive to support this judgement and sufficiently well documented that the reasoning can be critically assessed by similarly knowledgeable experts.

Best current judgements

When is the probability of an individual scientifically valuable?

[1] This individual is knowledgeable in the area

[2] the analysis that has led to this judgement has been sufficiently careful, thorough and exhaustive to support this judgement and sufficiently well documented that the reasoning can be critically assessed by similarly knowledgeable experts.

If a problem is important enough that the uncertainty analysis will have a large scientific, commercial or public policy implications, then best current judgements set a meaningful, rigorous standard for the analysis.

So, a worthwhile objective of an analysis is to produce the “best” current judgements of a specified expert (or group), in a transparent form.

Best current judgements

When is the probability of an individual scientifically valuable?

[1] This individual is knowledgeable in the area

[2] the analysis that has led to this judgement has been sufficiently careful, thorough and exhaustive to support this judgement and sufficiently well documented that the reasoning can be critically assessed by similarly knowledgeable experts.

If a problem is important enough that the uncertainty analysis will have a large scientific, commercial or public policy implications, then best current judgements set a meaningful, rigorous standard for the analysis.

So, a worthwhile objective of an analysis is to produce the "best" current judgements of a specified expert (or group), in a transparent form.

This is the objective of our analysis in the same way as the objective of a climate modeller is to represent actual climate as closely as possible.

To avoid the modeller's fallacy, we must be honest as to how well we achieve this aim. So, we need a way to express this.

Bayesian uncertainty quantification

Bayesian analysis is built on the collection of all your prior judgements, likelihood assessments etc. Inevitably, these involve many simplifications.

Bayesian uncertainty quantification

Bayesian analysis is built on the collection of all your prior judgements, likelihood assessments etc. Inevitably, these involve many simplifications.

Bayesian uncertainty quantification is by use of Bayes theorem.

Does Bayes theorem work (i.e. quantify real-world uncertainty) or is this a modelling simplification?

Bayesian uncertainty quantification

Bayesian analysis is built on the collection of all your prior judgements, likelihood assessments etc. Inevitably, these involve many simplifications.

Bayesian uncertainty quantification is by use of Bayes theorem.

Does Bayes theorem work (i.e. quantify real-world uncertainty) or is this a modelling simplification?

If A and B are both events, what does $\mathrm{P}(B|A)$ mean?

Bayesian uncertainty quantification

Bayesian analysis is built on the collection of all your prior judgements, likelihood assessments etc. Inevitably, these involve many simplifications.

Bayesian uncertainty quantification is by use of Bayes theorem.

Does Bayes theorem work (i.e. quantify real-world uncertainty) or is this a modelling simplification?

If A and B are both events, what does $\mathrm{P}(B|A)$ mean?

$\mathrm{P}(B)$ is your betting rate on B (e.g. your fair price for a ticket that pays 1 if B occurs, and pays 0 otherwise).

Bayesian uncertainty quantification

Bayesian analysis is built on the collection of all your prior judgements, likelihood assessments etc. Inevitably, these involve many simplifications.

Bayesian uncertainty quantification is by use of Bayes theorem.

Does Bayes theorem work (i.e. quantify real-world uncertainty) or is this a modelling simplification?

If A and B are both events, what does $\mathrm{P}(B|A)$ mean?

$\mathrm{P}(B)$ is your betting rate on B (e.g. your fair price for a ticket that pays 1 if B occurs, and pays 0 otherwise).

$\mathrm{P}(B|A)$ is your "called off" betting rate on B (e.g. your fair price for a ticket that pays 1 if B occurs, and pays 0 otherwise, if A occurs. If A doesn't occur your price is refunded).

Bayesian uncertainty quantification

Bayesian analysis is built on the collection of all your prior judgements, likelihood assessments etc. Inevitably, these involve many simplifications.

Bayesian uncertainty quantification is by use of Bayes theorem.

Does Bayes theorem work (i.e. quantify real-world uncertainty) or is this a modelling simplification?

If A and B are both events, what does $\mathrm{P}(B|A)$ mean?

$\mathrm{P}(B)$ is your betting rate on B (e.g. your fair price for a ticket that pays 1 if B occurs, and pays 0 otherwise).

$\mathrm{P}(B|A)$ is your “called off” betting rate on B (e.g. your fair price for a ticket that pays 1 if B occurs, and pays 0 otherwise, if A occurs. If A doesn’t occur your price is refunded).

This is NOT the same as the posterior probability that you will have for B if you find out that A occurs. There is no obvious relationship between the called off bet and posterior judgment at all. Equating the two meanings is an example of the modeller’s fallacy.

Bayesian uncertainty quantification

Bayesian analysis is built on the collection of all your prior judgements, likelihood assessments etc. Inevitably, these involve many simplifications.

Bayesian uncertainty quantification is by use of Bayes theorem.

Does Bayes theorem work (i.e. quantify real-world uncertainty) or is this a modelling simplification?

If A and B are both events, what does $\mathrm{P}(B|A)$ mean?

$\mathrm{P}(B)$ is your betting rate on B (e.g. your fair price for a ticket that pays 1 if B occurs, and pays 0 otherwise).

$\mathrm{P}(B|A)$ is your "called off" betting rate on B (e.g. your fair price for a ticket that pays 1 if B occurs, and pays 0 otherwise, if A occurs. If A doesn't occur your price is refunded).

This is NOT the same as the posterior probability that you will have for B if you find out that A occurs. There is no obvious relationship between the called off bet and posterior judgment at all. Equating the two meanings is an example of the modeller's fallacy.

Even worse: B often is the 'parameters' of some statistical model. The model may have been discarded when A is observed - so B no longer exists.

Expectation as a primitive

The Bayesian approach is hard for complicated problems because thinking about complicated problems is hard, and there is so much to think about.

Expectation as a primitive

The Bayesian approach is hard for complicated problems because thinking about complicated problems is hard, and there is so much to think about.

You can think about less if you use **EXPECTATION** rather than **PROBABILITY** as the primitive for expressing uncertainty judgments

That way you just make the expectation statements that you do need, (this might include some probability statements), rather than all of the uncountably many expectation statements that you might possibly need,

(see de Finetti "Theory of Probability", Wiley, 1974, who made expectation primitive for exactly that reason).

Expectation as a primitive

The Bayesian approach is hard for complicated problems because thinking about complicated problems is hard, and there is so much to think about.

You can think about less if you use **EXPECTATION** rather than **PROBABILITY** as the primitive for expressing uncertainty judgments

That way you just make the expectation statements that you do need, (this might include some probability statements), rather than all of the uncountably many expectation statements that you might possibly need,

(see de Finetti “Theory of Probability”, Wiley, 1974, who made expectation primitive for exactly that reason).

The version of Bayesian analysis which you get if you start with expectation is termed Bayes linear analysis, see

Bayes linear Statistics: Theory and Methods, 2007, (Wiley)

Michael Goldstein and David Wooff

Expectation as a primitive

Durham University

The Bayesian approach is hard for complicated problems because thinking about complicated problems is hard, and there is so much to think about.

You can think about less if you use **EXPECTATION** rather than **PROBABILITY** as the primitive for expressing uncertainty judgments

That way you just make the expectation statements that you do need, (this might include some probability statements), rather than all of the uncountably many expectation statements that you might possibly need,

(see de Finetti "Theory of Probability", Wiley, 1974, who made expectation primitive for exactly that reason).

The version of Bayesian analysis which you get if you start with expectation is termed Bayes linear analysis, see

Bayes linear Statistics: Theory and Methods, 2007, (Wiley)

Michael Goldstein and David Wooff

Our account of the meaning of Bayesian analysis requires expectation as primitive. (I know no such account with probability as primitive.)

Adjusted expectation

We treat expectation as primitive, follow the development of de Finetti, and define the expectation of a random quantity, Z as the value $\bar{z}$ that you would choose for z, if faced with the penalty

$$L = k(Z - z)^2,$$

where k is a constant defining the units of loss, and the penalty is paid in probability currency. [You can trade proper scoring rules for practical elicitation.]

Adjusted expectation

We treat expectation as primitive, follow the development of de Finetti, and define the expectation of a random quantity, Z as the value $\bar{z}$ that you would choose for z, if faced with the penalty

$$L = k(Z - z)^2,$$

where k is a constant defining the units of loss, and the penalty is paid in probability currency. [You can trade proper scoring rules for practical elicitation.]

The *adjusted* or *Bayes linear* expectation for B given $\boldsymbol{D}$, where $\boldsymbol{D} = (D_0, D_1, ..., D_s)$, with $D_0 = 1$ is the linear combination $\bar{\boldsymbol{a}}^T \boldsymbol{D}$ where $\bar{\boldsymbol{a}}$ is the value of $\boldsymbol{a}$ that you would choose if faced with the penalty

$$L = (B - \boldsymbol{a}^T \boldsymbol{D})^2$$

Adjusted expectation

We treat expectation as primitive, follow the development of de Finetti, and define the expectation of a random quantity, Z as the value $\bar{z}$ that you would choose for z, if faced with the penalty

$$L = k(Z - z)^2,$$

where k is a constant defining the units of loss, and the penalty is paid in probability currency. [You can trade proper scoring rules for practical elicitation.]

The *adjusted* or *Bayes linear* expectation for B given $\boldsymbol{D}$, where $\boldsymbol{D} = (D_0, D_1, ..., D_s)$, with $D_0 = 1$ is the linear combination $\bar{\boldsymbol{a}}^T \boldsymbol{D}$ where $\bar{\boldsymbol{a}}$ is the value of $\boldsymbol{a}$ that you would choose if faced with the penalty

$$L = (B - \boldsymbol{a}^T \boldsymbol{D})^2$$

It is given by

$$\mathrm{E}_{\boldsymbol{D}}(B) = \mathrm{E}(B) + \mathrm{Cov}(B, \boldsymbol{D})(\mathrm{Var}(\boldsymbol{D}))^{-1}(\boldsymbol{D} - \mathrm{E}(\boldsymbol{D}))$$

[Variances, covariances specified directly as primitive - or found by analysis.]

Belief adjustment and conditioning

Adjusted expectation is equivalent to conditional expectation in the particular case where $\boldsymbol{D}$ comprises the indicator functions for the elements of a partition, i.e. where each D_i takes value one or zero and precisely one element D_i will equal one, eg, if B is the indicator for an event, then

$$\mathrm{E}_{\boldsymbol{D}}(B) = \sum_i \mathrm{P}(B|D_i)D_i$$

Therefore, any interpretation of the meaning of belief adjustment immediately applies to full Bayes analysis.

Belief adjustment and conditioning

Adjusted expectation is equivalent to conditional expectation in the particular case where $\boldsymbol{D}$ comprises the indicator functions for the elements of a partition, i.e. where each D_i takes value one or zero and precisely one element D_i will equal one, eg, if B is the indicator for an event, then

$$\mathrm{E}_{\boldsymbol{D}}(B) = \sum_i \mathrm{P}(B|D_i)D_i$$

Therefore, any interpretation of the meaning of belief adjustment immediately applies to full Bayes analysis.

In order to establish links between our judgments now (conditional) and future (posterior), we need a meaningful notion of ‘temporal rationality’.

Belief adjustment and conditioning

Adjusted expectation is equivalent to conditional expectation in the particular case where $\boldsymbol{D}$ comprises the indicator functions for the elements of a partition, i.e. where each D_i takes value one or zero and precisely one element D_i will equal one, eg, if B is the indicator for an event, then

$$\mathrm{E}_{\boldsymbol{D}}(B) = \sum_i \mathrm{P}(B|D_i)D_i$$

Therefore, any interpretation of the meaning of belief adjustment immediately applies to full Bayes analysis.

In order to establish links between our judgments now (conditional) and future (posterior), we need a meaningful notion of 'temporal rationality'.

Our description is operational. It concerns preferences between random penalties, as assessed at different time points, considered as small cash penalties [or (better) payoffs in probability currency (i.e. tickets in a lottery with a single prize)].

Constraints on temporal preference

Current preferences, even when constrained by current conditional preferences given possible future outcomes, cannot require you to hold certain future preferences; for example, you may obtain further, hitherto unsuspected, information or insights into the problem before you come to make your future judgments.

Constraints on temporal preference

Current preferences, even when constrained by current conditional preferences given possible future outcomes, cannot require you to hold certain future preferences; for example, you may obtain further, hitherto unsuspected, information or insights into the problem before you come to make your future judgments.

It is much more compelling to suggest that future preferences may determine prior preferences. Suppose that you must choose between two random penalties, J and K.

Constraints on temporal preference

Current preferences, even when constrained by current conditional preferences given possible future outcomes, cannot require you to hold certain future preferences; for example, you may obtain further, hitherto unsuspected, information or insights into the problem before you come to make your future judgments.

It is much more compelling to suggest that future preferences may determine prior preferences. Suppose that you must choose between two random penalties, J and K.

For your future preferences to influence your current preferences, you must know what your future preference will be. You have a **sure preference** for J over K at (future) time t, if you know now, as a matter of logic, that at time t you will not express a strict preference for penalty K over penalty J.

Constraints on temporal preference

Current preferences, even when constrained by current conditional preferences given possible future outcomes, cannot require you to hold certain future preferences; for example, you may obtain further, hitherto unsuspected, information or insights into the problem before you come to make your future judgments.

It is much more compelling to suggest that future preferences may determine prior preferences. Suppose that you must choose between two random penalties, J and K.

For your future preferences to influence your current preferences, you must know what your future preference will be. You have a **sure preference** for J over K at (future) time t, if you know now, as a matter of logic, that at time t you will not express a strict preference for penalty K over penalty J.

Our (extremely weak) temporal consistency principle is that future sure preferences are respected by preferences today. We call this

The temporal sure preference principle *Suppose that you have a sure preference for J over K at (future) time t. Then you should not have a strict preference for K over J now.*

Adjusted and posterior expectation

For a particular random quantity Z, you specify a current expectation $\mathrm{E}(Z)$ and you intend to express a revised expectation $\mathrm{E}_t(Z)$ at time t. As $\mathrm{E}_t(Z)$ is unknown to you, you may express beliefs about this quantity.

Adjusted and posterior expectation

For a particular random quantity Z, you specify a current expectation $\mathrm{E}(Z)$ and you intend to express a revised expectation $\mathrm{E}_t(Z)$ at time t. As $\mathrm{E}_t(Z)$ is unknown to you, you may express beliefs about this quantity.

What does adjusted expectation $\mathrm{E}_{\boldsymbol{D}}(B)$ imply about the posterior assessment $\mathrm{E}_t(B)$ that we may make having observed $\boldsymbol{D}$?

Adjusted and posterior expectation

For a particular random quantity Z, you specify a current expectation $\mathrm{E}(Z)$ and you intend to express a revised expectation $\mathrm{E}_t(Z)$ at time t. As $\mathrm{E}_t(Z)$ is unknown to you, you may express beliefs about this quantity.

What does adjusted expectation $\mathrm{E}_{\boldsymbol{D}}(B)$ imply about the posterior assessment $\mathrm{E}_t(B)$ that we may make having observed $\boldsymbol{D}$?

The temporal sure preference principle implies that your actual posterior expectation, $\mathrm{E}_T(B)$, at time T when you have observed $\boldsymbol{D}$, satisfies

Adjusted and posterior expectation

For a particular random quantity Z, you specify a current expectation $\mathrm{E}(Z)$ and you intend to express a revised expectation $\mathrm{E}_t(Z)$ at time t. As $\mathrm{E}_t(Z)$ is unknown to you, you may express beliefs about this quantity.

What does adjusted expectation $\mathrm{E}_{\boldsymbol{D}}(B)$ imply about the posterior assessment $\mathrm{E}_t(B)$ that we may make having observed $\boldsymbol{D}$?

The temporal sure preference principle implies that your actual posterior expectation, $\mathrm{E}_T(B)$, at time T when you have observed $\boldsymbol{D}$, satisfies

$$B = \mathrm{E}_T(B) \oplus S$$

Adjusted and posterior expectation

For a particular random quantity Z, you specify a current expectation $\mathrm{E}(Z)$ and you intend to express a revised expectation $\mathrm{E}_t(Z)$ at time t. As $\mathrm{E}_t(Z)$ is unknown to you, you may express beliefs about this quantity.

What does adjusted expectation $\mathrm{E}_{\boldsymbol{D}}(B)$ imply about the posterior assessment $\mathrm{E}_t(B)$ that we may make having observed $\boldsymbol{D}$?

The temporal sure preference principle implies that your actual posterior expectation, $\mathrm{E}_T(B)$, at time T when you have observed $\boldsymbol{D}$, satisfies

$$B = \mathrm{E}_T(B) \oplus S$$

$$\mathrm{E}_T(B) = \mathrm{E}_{\boldsymbol{D}}(B) \oplus R$$

Adjusted and posterior expectation

For a particular random quantity Z, you specify a current expectation $\mathrm{E}(Z)$ and you intend to express a revised expectation $\mathrm{E}_t(Z)$ at time t. As $\mathrm{E}_t(Z)$ is unknown to you, you may express beliefs about this quantity.

What does adjusted expectation $\mathrm{E}_{\boldsymbol{D}}(B)$ imply about the posterior assessment $\mathrm{E}_t(B)$ that we may make having observed $\boldsymbol{D}$?

The temporal sure preference principle implies that your actual posterior expectation, $\mathrm{E}_T(B)$, at time T when you have observed $\boldsymbol{D}$, satisfies

$$B = \mathrm{E}_T(B) \oplus S$$

$$\mathrm{E}_T(B) = \mathrm{E}_{\boldsymbol{D}}(B) \oplus R$$

where S, R each have, a priori, zero expectation and are uncorrelated with each other and with $\boldsymbol{D}$.

Adjusted and posterior expectation

For a particular random quantity Z, you specify a current expectation $\mathrm{E}(Z)$ and you intend to express a revised expectation $\mathrm{E}_t(Z)$ at time t. As $\mathrm{E}_t(Z)$ is unknown to you, you may express beliefs about this quantity.

What does adjusted expectation $\mathrm{E}_{\boldsymbol{D}}(B)$ imply about the posterior assessment $\mathrm{E}_t(B)$ that we may make having observed $\boldsymbol{D}$?

The temporal sure preference principle implies that your actual posterior expectation, $\mathrm{E}_T(B)$, at time T when you have observed $\boldsymbol{D}$, satisfies

$$B = \mathrm{E}_T(B) \oplus S$$

$$\mathrm{E}_T(B) = \mathrm{E}_{\boldsymbol{D}}(B) \oplus R$$

where S, R each have, a priori, zero expectation and are uncorrelated with each other and with $\boldsymbol{D}$.

Therefore, $\mathrm{E}_{\boldsymbol{D}}(B)$ resolves some of your current uncertainty for $\mathrm{E}_T(B)$ which resolves some of your uncertainty for B.

[Actual amount of variance resolved is $\mathrm{Cov}(B, \boldsymbol{D})(\mathrm{Var}(\boldsymbol{D}))^{-1}\mathrm{Cov}(\boldsymbol{D}, B)$]

Why this works

You can make a current expectation for $(Z - \mathrm{E}_t(Z))^2$.

Why this works

You can make a current expectation for $(Z - \mathrm{E}_t(Z))^2$.

Suppose that F is any random quantity whose value you will surely know by time t. Suppose that you assess a current expectation for $(Z - F)^2$.

Why this works

You can make a current expectation for $(Z - \mathrm{E}_t(Z))^2$.
Suppose that F is any random quantity whose value you will surely know by time $t.$ Suppose that you assess a current expectation for $(Z - F)^2$.
To satisfy temporal sure preference you must now assign

$$\mathrm{E}((Z - \mathrm{E}_t(Z))^2) \leq \mathrm{E}((Z - F)^2)$$

Why this works

You can make a current expectation for $(Z - \mathrm{E}_t(Z))^2$.

Suppose that F is any random quantity whose value you will surely know by time t. Suppose that you assess a current expectation for $(Z - F)^2$.

To satisfy temporal sure preference you must now assign

$$\mathrm{E}((Z - \mathrm{E}_t(Z))^2) \leq \mathrm{E}((Z - F)^2)$$

For any set of random quantities $X = (X_1, X_2, \ldots)$, create the inner product space $I(X)$ whose vectors are linear combinations of the elements of X, with covariance as the inner product.

Why this works

You can make a current expectation for $(Z - \mathrm{E}_t(Z))^2$.

Suppose that F is any random quantity whose value you will surely know by time t. Suppose that you assess a current expectation for $(Z - F)^2$.

To satisfy temporal sure preference you must now assign

$$\mathrm{E}((Z - \mathrm{E}_t(Z))^2) \leq \mathrm{E}((Z - F)^2)$$

For any set of random quantities $X = (X_1, X_2, ...)$, create the inner product space $I(X)$ whose vectors are linear combinations of the elements of X, with covariance as the inner product.

If D is a vector whose elements will surely be known by time t, then, for any quantity Y, then $\mathrm{E}_D(Y)$ is the orthogonal projection of Y into $I(D)$.

Why this works

You can make a current expectation for $(Z - \mathrm{E}_t(Z))^2$.
Suppose that F is any random quantity whose value you will surely know by time t. Suppose that you assess a current expectation for $(Z - F)^2$.
To satisfy temporal sure preference you must now assign

$$\mathrm{E}((Z - \mathrm{E}_t(Z))^2) \leq \mathrm{E}((Z - F)^2)$$

For any set of random quantities $X = (X_1, X_2, \ldots)$, create the inner product space $I(X)$ whose vectors are linear combinations of the elements of X, with covariance as the inner product.
If D is a vector whose elements will surely be known by time t, then, for any quantity Y, then $\mathrm{E}_D(Y)$ is the orthogonal projection of Y into $I(D)$.
If we let $I(D, \mathrm{E}_t(Y))$ be the inner product space formed by adding $\mathrm{E}_t(Y)$ to $I(D)$, then $\mathrm{E}_t(Y)$ is the orthogonal projection of Y into $I(D, \mathrm{E}_t(Y))$.

Conditional Probabilities and Best expert judgements

Durham University

If $\boldsymbol{D}$ represents a partition, then conditional and posterior judgements relate by

$$\mathrm{E}_T(B) = \mathrm{E}(B|\boldsymbol{D}) + R$$

where

$$\mathrm{E}(\boldsymbol{R}|\boldsymbol{D_i}) = 0, \forall i$$

Conditional Probabilities and Best expert judgements

If $\boldsymbol{D}$ represents a partition, then conditional and posterior judgements relate by

$$\mathrm{E}_T(B) = \mathrm{E}(B|\boldsymbol{D}) + R$$

where

$$\mathrm{E}(\boldsymbol{R}|\boldsymbol{D_i}) = 0, \forall i$$

This relation holds whatever the posterior extension consistent with the current conditional specification. In particular, if we view the Bayes analysis as modelling best expert judgements for the problem, then the conditional Bayes analysis, as a model for such judgements, reduces, but does not eliminate, uncertainty about what those judgements should be.

Conditional Probabilities and Best expert judgements

If $\boldsymbol{D}$ represents a partition, then conditional and posterior judgements relate by

$$\mathrm{E}_T(B) = \mathrm{E}(B|\boldsymbol{D}) + R$$

where

$$\mathrm{E}(\boldsymbol{R}|\boldsymbol{D_i}) = 0, \forall i$$

This relation holds whatever the posterior extension consistent with the current conditional specification. In particular, if we view the Bayes analysis as modelling best expert judgements for the problem, then the conditional Bayes analysis, as a model for such judgements, reduces, but does not eliminate, uncertainty about what those judgements should be.

This is no different than any other relationship between a real quantity and a model for that quantity, except that, for probabilistic analysis, we can rigorously derive the corresponding relationship, under very weak, plausible and testable assumptions.

Updating judgements over models

If $\boldsymbol{X_1}, \boldsymbol{X_2}, ...$ are infinite Second Order Exchangeable (SOE), i.e. each has same mean, variance and all pairwise covariances the same, then
$\boldsymbol{X}_i = \boldsymbol{M} \oplus \boldsymbol{R}_i \ \forall i$ where $R_1, R_2, ...$ are SOE, uncorrelated, mean zero.

Updating judgements over models

If $\boldsymbol{X_1}, \boldsymbol{X_2}, ...$ are infinite Second Order Exchangeable (SOE), i.e. each has same mean, variance and all pairwise covariances the same, then $\boldsymbol{X}_i = \boldsymbol{M} \oplus \boldsymbol{R}_i \ \forall i$ where $R_1, R_2, ...$ are SOE, uncorrelated, mean zero. Suppose you will observe a sample $(\boldsymbol{X_{[n]}} = \boldsymbol{X_1}, ..., \boldsymbol{X_n})$, by time T. You don't know whether $\boldsymbol{X_{n+1}}, \boldsymbol{X_{n+2}}, ...$ will be SOE at time T.

Updating judgements over models

If $\boldsymbol{X_1}, \boldsymbol{X_2}, ...$ are infinite Second Order Exchangeable (SOE), i.e. each has same mean, variance and all pairwise covariances the same, then $\boldsymbol{X}_i = \boldsymbol{M} \oplus \boldsymbol{R}_i \; \forall i$ where $R_1, R_2, ...$ are SOE, uncorrelated, mean zero.

Suppose you will observe a sample $(\boldsymbol{X}_{[\boldsymbol{n}]} = \boldsymbol{X_1}, ..., \boldsymbol{X_n})$, by time T. You don't know whether $\boldsymbol{X_{n+1}}, \boldsymbol{X_{n+2}}, ...$ will be SOE at time T.

Theorem Given (i) Temporal Sure Preference

(ii) current judgement that sequence $\mathrm{E}_T(\boldsymbol{X_{n+1}}), \mathrm{E}_T(\boldsymbol{X_{n+2}}), ...$ is SOE,

Updating judgements over models

If $\boldsymbol{X_1}, \boldsymbol{X_2}, ...$ are infinite Second Order Exchangeable (SOE), i.e. each has same mean, variance and all pairwise covariances the same, then
$\boldsymbol{X}_i = \boldsymbol{M} \oplus \boldsymbol{R}_i \ \forall i$ where $R_1, R_2, ...$ are SOE, uncorrelated, mean zero.
Suppose you will observe a sample $(\boldsymbol{X_{[n]}} = \boldsymbol{X_1}, ..., \boldsymbol{X_n})$, by time T. You don't know whether $\boldsymbol{X_{n+1}}, \boldsymbol{X_{n+2}}, ...$ will be SOE at time T.
Theorem Given (i) Temporal Sure Preference
(ii) current judgement that sequence $\mathrm{E}_T(\boldsymbol{X_{n+1}}), \mathrm{E}_T(\boldsymbol{X_{n+2}}), ...$ is SOE, you can construct a further quantity, $\mathrm{E}_T(\boldsymbol{M})$, which decomposes your judgements about any future outcome $\boldsymbol{X_j} = \boldsymbol{M} \oplus \boldsymbol{R}_j, \ j > n$ as

Updating judgements over models

If $\boldsymbol{X_1}, \boldsymbol{X_2}, ...$ are infinite Second Order Exchangeable (SOE), i.e. each has same mean, variance and all pairwise covariances the same, then $\boldsymbol{X}_i = \boldsymbol{M} \oplus \boldsymbol{R}_i \; \forall i$ where $R_1, R_2, ...$ are SOE, uncorrelated, mean zero. Suppose you will observe a sample $(\boldsymbol{X_{[n]}} = \boldsymbol{X_1}, ..., \boldsymbol{X_n})$, by time T. You don't know whether $\boldsymbol{X_{n+1}}, \boldsymbol{X_{n+2}}, ...$ will be SOE at time T.

Theorem Given (i) Temporal Sure Preference

(ii) current judgement that sequence $\mathrm{E}_T(\boldsymbol{X_{n+1}}), \mathrm{E}_T(\boldsymbol{X_{n+2}}), ...$ is SOE, you can construct a further quantity, $\mathrm{E}_T(\boldsymbol{M})$, which decomposes your judgements about any future outcome $\boldsymbol{X_j} = \boldsymbol{M} \oplus \boldsymbol{R}_j,\; j > n$ as

$$\begin{aligned}
\boldsymbol{X_j} - \mathrm{E}(\boldsymbol{X}) \quad = \quad & [\boldsymbol{M} - \mathrm{E}_T(\boldsymbol{M})] \\
& \oplus [\mathrm{E}_T(\boldsymbol{M}) - \mathrm{E}_{\boldsymbol{X_{[n]}}}(\boldsymbol{M})] \\
& \oplus [\mathrm{E}_{\boldsymbol{X_{[n]}}}(\boldsymbol{M}) - \mathrm{E}(\boldsymbol{M})] \\
& \oplus [\boldsymbol{R_j} - \mathrm{E}_T(\boldsymbol{R_j})] \\
& \oplus [\mathrm{E}_T(\boldsymbol{R_j})]
\end{aligned}$$

Does Bayes theorem work?

Does Bayes analysis tell you how to update your beliefs given data?

No. That's the modeller's fallacy.

Does Bayes theorem work?

Does Bayes analysis tell you how to update your beliefs given data?
No. That's the modeller's fallacy.

Is Bayes analysis informative for belief updating?
Yes. Conditional probabilities, adjusted expectations, etc. resolve some of the uncertainty about such updating, in a precise and well-defined way.

Does Bayes theorem work?

Does Bayes analysis tell you how to update your beliefs given data?
No. That's the modeller's fallacy.

Is Bayes analysis informative for belief updating?
Yes. Conditional probabilities, adjusted expectations, etc. resolve some of the uncertainty about such updating, in a precise and well-defined way.

The notion of Best Current Judgements, and how informative our analysis is for them, is a useful and constructive way to give practical meaning to a Bayesian analysis. There is a logical structure to help us to do this.

Some References

Goldstein, M. (1994) Revising exchangeable beliefs: subjectivist foundations for the inductive argument, in Aspects of Uncertainty, a tribute to D.V. Lindley , 201-222

Goldstein, M (1997) Prior inferences for posterior judgements (1997), in *Structures and norms in Science*, M.C.D. Chiara et. al. eds., Kluwer, 55-71.

Goldstein, M. (2006) Subjective Bayesian analysis: principles and practice Bayesian Analysis 1 403-420

Goldstein, M. and Rougier, J.C. (2009) Reified Bayesian modelling and inference for physical systems, JSPI139, 1221-1239

Goldstein, M. (2011). External Bayesian analysis for computer simulators. In BAYESIAN STATISTICS 9. Bernardo, J.M. et al, eds, Oxford University Press.

And check out the website for the

Managing Uncertainty in Complex Models (MUCM) project

[A consortium of Aston, Durham, LSE, Sheffield and Southampton all hard at work on developing technology for computer model uncertainty problems.]