

Deciding When Evidence Is Incomplete

SRAI Executive Pathway Module 3 Canonical Executive Lesson

Executive purpose. This module equips accountable leaders to update beliefs when new evidence arrives, translate posterior uncertainty into decisions under asymmetric loss, and determine whether to act, pilot, wait and learn, or escalate. Its central discipline is simple but demanding: a probability is not a decision, and a model is not an authorization.

Control	Specification
Module	EP-M03 Deciding When Evidence Is Incomplete
Asset	EP-M03-A01 Canonical Executive Lesson
Version	v0.1
Status	G2 Content Prototype
Primary source	Former Book 6 Chapter 3 Bayesian Decision Analysis
Companion	EP-M03-A08 Companion Notebook v0.2
Terminal artifact	Evidence Update and Decision Loss Matrix
Decision vocabulary	ACT, PILOT, WAIT AND LEARN, or ESCALATE
Priority sectors	Agriculture, Education, Health, and Habitat
Date	18 September 2026

Main conclusion. Evidence changes rational action only through its effect on beliefs and through the consequences attached to those beliefs. Strong governance therefore requires an explicit prior, a defensible likelihood model, a transparent loss structure, sensitivity analysis, and a named human authority.

1 Executive Learning Contract

The decision problem

Executives rarely choose with complete evidence. They approve a programme before outcomes are fully observed, allocate scarce funds while causal claims remain contested, respond to an emerging hazard before its severity is known, and decide whether another study is worth the delay it imposes. The discipline developed here is not a search for certainty. It is a governed method for making uncertainty explicit and deciding what should happen next.

Learning outcomes

- Represent uncertain states, actions, evidence, consequences, and authority as one decision system.
- Use Bayes rule and likelihood ratios to update beliefs without confusing evidence with causation.
- Select a Bayes action by minimizing posterior expected loss rather than maximizing probability alone.
- Test whether the recommendation reverses under credible priors, likelihoods, losses, missing states, or distributional constraints.
- Estimate the value of perfect and imperfect information and compare it with monetary cost, delay, and foregone benefit.
- Produce an Evidence Update and Decision Loss Matrix that records the decision, its assumptions, monitoring conditions, and accountable authority.

Boundary with adjacent modules

Module	Primary question	What carries forward
Module 2	How should value, risk, utility, and hard constraints be made explicit?	Consequences, utilities, constraints, sensitivity, and information value.
Module 3	How should beliefs change with evidence, and what action follows under asymmetric loss?	Posterior beliefs, Bayes action, robustness, and governed learning.
Module 4	How should alternatives be compared across plural and competing objectives?	A defensible evidence and loss record for multi-criteria comparison.

Prerequisite. The learner should already distinguish outcomes from actions, understand expected value as a benchmark, and recognize that rights and safety

constraints may be non-compensatory. Module 3 does not replace those ideas; it makes uncertainty about the state of the world operational.

2 Opening Case Agriculture Under Emerging Evidence

A regional agriculture authority must decide whether to authorize an early crop-protection intervention. The season is unusually humid, field reports indicate a possible fungal outbreak, and delayed action could produce severe yield loss. Yet treatment has financial, ecological, and trust costs if the outbreak is limited. The authority cannot observe the true outbreak state directly before acting.

Decision element	Canonical case
States	Limited outbreak; severe outbreak
Actions	Monitor; intervene
Prior	$P(\text{limited}) = 0.60$; $P(\text{severe}) = 0.40$
Evidence	A positive surveillance signal
Likelihood	$P(\text{positive} \mid \text{limited}) = 0.30$; $P(\text{positive} \mid \text{severe}) = 0.80$
Loss if monitor	0 if limited; 20 if severe
Loss if intervene	8 if limited; 0 if severe
Authority	Named regional authorization body with agronomic, fiscal, and environmental duties

Why this is an executive decision

The calculation is short; the governance is not. The prior may blend historical incidence with local expertise. The likelihood may come from a surveillance model whose transportability is uncertain. Losses mix yield, cost, ecological harm, food security, and distributional exposure. A mathematically correct answer can therefore be institutionally wrong if the inputs are unowned, the decision boundary is hidden, or a serious state has been omitted.

The executive task is to keep four statements separate: what was believed before the signal, how compatible the signal is with each state, what each action would cost

under each state, and who has authority to accept the residual risk. This separation prevents persuasive evidence language from silently determining policy.

3 The Anatomy of a Governed Decision

A decision under incomplete evidence contains at least five objects. A state is a condition that matters but is not fully observed. An action is a controllable choice. Evidence is an observation whose probability depends on the state. A consequence is what follows from pairing an action with a state. Authority is the legitimate capacity to approve, revise, defer, or stop the action.

Object	Executive question	Governance test
State	What could be true?	Are material states explicit, mutually intelligible, and sufficiently complete?
Action	What can we choose now?	Are inaction, pilot, staged action, and escalation represented where feasible?
Evidence	What was observed?	Is provenance, timing, measurement error, and dependence declared?
Consequence	What is at stake?	Are losses symmetric or asymmetric, and who bears them?
Authority	Who may decide?	Is the decision within mandate, rights, safety, and budget constraints?

Decision quality before outcome quality

A good decision can produce a bad outcome because uncertainty remains. A bad decision can occasionally produce a good outcome by luck. Governance should therefore audit the quality of the process at the time of authorization: Were the states and actions complete enough? Were evidence and losses defensible? Was sensitivity tested? Were monitoring and review conditions specified? Outcome review remains essential, but retrospective success cannot repair an incoherent decision record.

The decision record

The Evidence Update and Decision Loss Matrix is the terminal artifact of this module. It is not merely a calculation sheet. It links assumptions, evidence, losses, reversals, information value, monitoring, and authorization so that another qualified reviewer can reproduce the recommendation and see why a different recommendation might become justified later.

4 Beliefs Before Evidence

A prior distribution records uncertainty before the focal evidence is incorporated. It may be empirical, expert-elicited, model-based, or a structured combination. Calling a prior subjective does not make it arbitrary; calling it data-driven does not make it neutral. The relevant question is whether its provenance, population, period, exclusions, and uncertainty are fit for the decision.

Base rates and prior provenance

Base rates discipline interpretation. A highly sensitive signal can still have modest positive predictive value when the severe state is rare. Conversely, a weak signal may matter when the prior risk is already high. Executives should insist that the baseline population matches the decision population and that shifts in climate, implementation capacity, demography, or data-generating processes are considered.

Prior source	Strength	Risk	Required disclosure
Historical frequency	Observable and auditable	Past may not represent current regime	Period, denominator, missingness, structural change
Expert elicitation	Uses contextual knowledge	Overconfidence, anchoring, group effects	Experts, protocol, disagreement, calibration
Predictive model	Scalable and internally consistent	Dataset shift and hidden assumptions	Training population, validation, uncertainty
Reference class	Supports comparison	Reference-class selection can be strategic	Inclusion rule and relevance to current case
Mixture or range	Represents disagreement	Can hide incompatible causal stories	Weights, rationale, and decision sensitivity

Represent disagreement rather than averaging it away

If credible groups disagree about the prior, record the range or mixture and test the decision under each view. A consensus average may imply a belief held by nobody and can obscure whether the recommendation depends on a contested assumption. Disagreement is decision-relevant when it changes the preferred action, the value of further evidence, or the acceptable review interval.

5 Evidence Likelihoods and Diagnostic Weight

A likelihood asks how probable the observed evidence would be if a particular state were true. It is not the probability that the state is true. Confusing $P(\text{evidence} \mid \text{state})$ with $P(\text{state} \mid \text{evidence})$ is the inverse-probability error and is one of the most consequential mistakes in risk communication.

$$\text{Likelihood ratio} = P(\text{evidence} \mid \text{severe state}) / P(\text{evidence} \mid \text{limited state})$$

For the agriculture signal, the positive likelihood ratio is $0.80 / 0.30 = 2.67$. The evidence is therefore about 2.67 times as compatible with a severe outbreak as with a limited one. This is meaningful evidence, but it does not determine the posterior without the prior.

Evidence quality questions

- Measurement validity: Does the signal measure the claimed condition or a correlated proxy?
- Calibration: Across comparable cases, do stated probabilities correspond to observed frequencies?
- Transportability: Does validation in one geography, institution, or subgroup extend to the present setting?
- Conditional dependence: Are multiple signals genuinely independent once the state is known?
- Selection: Was the evidence produced only for cases already considered high risk?
- Timing: Does the evidence describe the state at the decision time, or an earlier and potentially obsolete condition?
- Adversarial incentives: Could reporting, coding, or observation change because actors benefit from a classification?

Do not double count correlated evidence

Sequential updating is valid when each likelihood is conditioned on all information already incorporated. Two reports derived from the same sensor network, source document, or model are not two independent confirmations. Treating them as

independent can produce extreme posteriors without new information. The evidence ledger should identify common sources and explicit dependence assumptions.

6 Bayesian Updating

Bayes rule combines the prior with the likelihood and renormalizes across states. It does not manufacture certainty. The posterior remains conditional on the declared states, model, evidence, and assumptions.

$$P(\text{state} \mid \text{evidence}) = P(\text{evidence} \mid \text{state}) \times P(\text{state}) / P(\text{evidence})$$

$$\text{Posterior odds} = \text{Prior odds} \times \text{Likelihood ratio}$$

Canonical calculation

State	Prior	Positive likelihood	Unnormalized weight	Posterior
Limited	0.60	0.30	0.18	0.18 / 0.50 = 0.36
Severe	0.40	0.80	0.32	0.32 / 0.50 = 0.64

The positive signal changes the severe-outbreak probability from 0.40 to 0.64. The posterior does not mean that 64 percent of the crop will be lost, that the signal is 64 percent accurate, or that intervention is automatically authorized. It means that, within the stated two-state model and inputs, the severe state has posterior probability 0.64.

Sequential updating

When a second signal arrives, the first posterior can become the next prior, provided the second likelihood is correctly conditioned. In practice, preserve the evidence sequence, timestamps, dependence assumptions, and model version. An audit trail should make it possible to reconstruct what was known at each decision point, not only the final posterior.

Posterior predictive reasoning

Executives often care less about the latent state itself than about future observable consequences. Posterior predictive distributions propagate state uncertainty into yields, hospital demand, learning outcomes, construction failure, or other operational quantities. They are useful for planning, but their quality still depends on the model linking state to outcome.

7 A Posterior Is Not a Decision

Prediction asks what is likely to be true. Decision analysis asks what should be done given beliefs, consequences, constraints, and authority. The most probable state need not determine the optimal action. A lower-probability state can dominate when failing to respond carries very large or irreversible loss.

Question	Analytical object	Common error
What state is most likely?	Posterior probability	Treating classification as authorization
What action minimizes expected harm?	Posterior expected loss	Ignoring false-positive, false-negative, or delay losses
Is the action permissible?	Rights, safety, legal, fiscal, and institutional constraints	Allowing expected value to override a hard boundary
Who decides?	Named accountable authority	Delegating authority to a model or analyst

Predictive accuracy versus decision value

A model can improve accuracy while leaving the preferred action unchanged; such improvement may have little decision value. A modestly accurate signal can have high decision value if it moves cases across a consequential action threshold. Model evaluation should therefore include calibration and discrimination, but also decision-specific loss, subgroup impact, stability, and information value.

SRAI Principle 143

Posterior beliefs inform but do not authorize action. Decisions require explicit losses, sensitivity analysis, and accountable human authority.

This principle is a governance boundary. It does not require executives to ignore quantitative recommendations. It requires that a recommendation remain contestable, reproducible, and subordinate to legitimate decision rights.

8 Loss Functions and Asymmetric Consequences

A loss function assigns a consequence score to each action-state pair. Loss may be expressed in money, utility units, health burden, service failure, learning loss,

ecological damage, or a structured composite. What matters is that lower values are preferred and that the scale is used consistently within the decision model.

Action	Limited state	Severe state	Interpretation
Monitor	0	20	Avoids unnecessary treatment but risks severe delay loss
Intervene	8	0	Incurs treatment cost if limited but avoids severe outbreak loss

Asymmetry changes the threshold

If the loss of monitoring during a severe outbreak greatly exceeds the loss of unnecessary intervention, the rational threshold for intervention can be well below 0.50. This is not irrational risk aversion; it is the direct implication of unequal consequences. Conversely, an intrusive or irreversible intervention may require a high posterior probability or a binding safety review.

Loss elicitation

- Define the perspective: organization, affected population, treasury, regulator, or multi-party governance body.
- Declare the time horizon and discounting rule; immediate cost and long-term harm are not interchangeable.
- Separate measurable consequences from ethical or legal constraints that should remain gates.
- Record distributional incidence: the same aggregate loss can conceal severe burden on a vulnerable subgroup.
- Use ranges when precision is not defensible and test whether the decision changes across them.
- Avoid reverse-engineering losses to justify a preferred action; elicit them before inspecting the recommendation where possible.

9 Bayes Actions and Posterior Expected Loss

Posterior expected loss(action) = sum over states of Loss(action, state) x P(state | evidence)

Bayes action = the feasible action with minimum posterior expected loss

Worked decision

Action	Calculation	Posterior expected loss
Monitor	$0 \times 0.36 + 20 \times 0.64$	12.80
Intervene	$8 \times 0.36 + 0 \times 0.64$	2.88

Intervene is the Bayes action because 2.88 is lower than 12.80. The expected-loss advantage is 9.92 units. That advantage is conditional on the two-state model, posterior, loss matrix, feasibility constraints, and decision horizon.

Decision threshold

Let p be the probability of a severe outbreak. Monitoring has expected loss $20p$. Intervention has expected loss $8(1 - p)$. The two actions are equally preferred when $20p = 8(1 - p)$, giving $p = 8/28 = 0.286$. Intervention is preferred whenever the severe-state probability exceeds about 28.6 percent.

The threshold exposes the role of loss asymmetry. It also gives executives a precise question: Is the credible posterior range entirely above, entirely below, or straddling 0.286? A recommendation is robust when all defensible inputs keep it on the same side of the threshold.

Feasibility and dominance

Expected loss should be calculated only for feasible actions. If an action violates a rights, safety, capacity, or legal constraint, it is not rescued by a favorable average. Before calculating, remove infeasible actions or represent the constraint explicitly. Also check dominance: an action that is no better in any state and worse in at least one should generally be excluded.

10 Sensitivity and Reversal Analysis

Sensitivity analysis asks whether the recommendation survives credible variation in the inputs. It is not a ritual of changing numbers at random. Each range should have an evidential or governance basis, and the result should identify the assumptions capable of reversing the action.

Input	Sensitivity question	Executive interpretation
Prior	Does the decision change across credible base rates?	If yes, prior provenance is a decision-critical issue.

Input	Sensitivity question	Executive interpretation
Likelihood	Does plausible measurement error reverse the posterior?	If yes, validate or diversify evidence.
Loss	Do stakeholder or horizon choices change the action threshold?	If yes, the value judgment requires authority.
Dependence	Does correlated evidence materially reduce effective information?	If yes, revise the evidence ledger.
State space	Does a missing severe or benign state change the action?	If yes, expand the model or escalate.

One-way and joint sensitivity

One-way sensitivity varies one input while holding others fixed and is useful for interpretation. Joint sensitivity varies several inputs together and is more realistic when assumptions are correlated. Scenario-based sensitivity can represent coherent worlds, such as a climate-shift scenario with higher outbreak prior and weaker historical transportability.

Decision regions

A decision-region plot maps which action is preferred across a grid of priors and losses. It converts a point recommendation into a stability map. Executives should focus on the distance between the current estimate and the reversal boundary, the plausibility of crossing it, and whether monitoring can detect movement before consequences become irreversible.

Robustness rather than false precision

When inputs are weak, a narrow posterior can be misleading. A robust action performs acceptably across plausible models, even if it is not optimal under any single point estimate. Robustness may justify a reversible pilot, a precautionary safeguard, or escalation when model disagreement is too consequential for ordinary authorization.

11 Likelihood and Model Uncertainty

Standard calculations often treat likelihoods as known. In real systems, sensitivity, specificity, reporting rates, and model parameters are estimated with error. The companion notebook represents this uncertainty by sampling plausible likelihoods and recomputing posteriors and decisions. The output is a distribution of recommendations, not merely a distribution of probabilities.

Three layers of uncertainty

Layer	Example	Appropriate response
State uncertainty	Severe or limited outbreak	Bayesian updating and expected loss
Parameter uncertainty	Unknown signal sensitivity	Posterior or simulation over parameter values
Model uncertainty	Omitted state or wrong causal structure	Alternative models, stress tests, external review, escalation

Parameter simulation should not be mistaken for model completeness. Thousands of draws from the wrong model can create an illusion of rigor. The analyst must justify the uncertainty distributions, preserve their provenance, and report the frequency with which each action is preferred.

Calibration and validation

Calibration tests whether predicted probabilities align with observed frequencies. Validation should be local enough to support the decision population, recent enough to reflect current conditions, and stratified enough to reveal subgroup failure. A model may be globally calibrated while systematically misestimating risk for a priority region or vulnerable population.

Model averaging and model sets

When several defensible models exist, decision analysis can weight them or evaluate each separately. Weighting is appropriate only when the weights themselves have a defensible basis. Otherwise, report the model set and whether the preferred action is invariant. If models imply materially different actions and the consequences are severe, escalation is a substantive result, not an analytical failure.

12 The Value of Information

Information has decision value only if it can change action or improve consequences. Better prediction is not automatically worth purchasing. The value of information must be compared with acquisition cost, delay, implementation capacity, and the harm incurred while waiting.

Expected value of perfect information

$$\text{EVPI} = \text{Current minimum expected loss} - \text{Expected loss with perfect state information}$$

In the canonical case, current minimum expected loss is 2.88. With perfect information, the authority would monitor in the limited state and intervene in the severe state, producing zero loss under the simplified matrix. EVPI is therefore 2.88 loss units. No information source can be worth more than this upper bound within the same model and loss scale.

Expected value of sample information

$$\text{EVSI} = \text{Current minimum expected loss} - \text{Expected loss under the optimal evidence-contingent policy}$$

The expanded notebook evaluates an imperfect test whose outcomes can change the chosen action. Under the declared test characteristics, EVSI is 1.168 loss units. This is a gross value. The test is worthwhile only if its monetary and non-monetary cost, delay loss, and implementation burden are lower than the value generated.

Information question	Decision rule
Can the result arrive before the decision closes?	If no, operational value may be zero even when statistical value is positive.
Can the result change action?	If no, EVSI is zero under the current decision model.
Does testing expose anyone to harm or inequity?	Include that consequence and apply relevant rights or safety gates.
Is the test repeatable and governed?	Record quality control, custody, interpretation, and authority.
What is the net value?	Proceed when gross information value exceeds cost and delay, subject to constraints.

13 Missing States and Model Misspecification

The most dangerous error may not be a wrong probability but a missing state. A two-state model can exclude catastrophic, adversarial, systemic, or compound conditions. Because posterior probabilities sum to one over the declared states, the model can appear complete while being substantively incomplete.

Stress test the state space

- Ask what would make both current actions inappropriate.
- Search for irreversible harm, legal breach, cascading failure, and vulnerable-population exposure.
- Consider common-cause events that invalidate assumed independence.
- Identify states suppressed by data collection, institutional incentives, or narrow disciplinary framing.
- Invite domain, community, operational, and adversarial perspectives before finalizing the state set.

Emergency state extension

The companion notebook adds an emergency state to demonstrate how a previously preferred action can become inadequate. If the extended model introduces a consequence beyond the authority's mandate or risk appetite, the appropriate result is ESCALATE. Escalation means the model has identified a governance boundary, not that analysis should continue until it produces a comfortable answer.

Trigger	Required response
Material state omitted	Expand the model and repeat the decision analysis.
No feasible action protects a hard constraint	Escalate to the competent authority.
Model set yields conflicting actions under severe consequences	Seek independent review or adopt a reversible safeguard.
Evidence-generating process has changed	Suspend reliance, revalidate, and shorten the review interval.

14 Distributional Effects Rights and Hard Constraints

Expected loss aggregates. Governance must disaggregate. Two actions can have the same total expected loss while imposing very different burdens across communities,

genders, income groups, regions, disability status, or generations. An executive record should show who benefits, who bears risk, and who has voice in the loss specification.

Distributional visibility

Question	Minimum evidence
Who bears false-positive loss?	Costs, restrictions, stigma, displacement, or ecological burden by group
Who bears false-negative loss?	Exposure to untreated hazard, service failure, or missed support by group
Who bears delay loss?	Time-sensitive harm and ability to absorb waiting
Who can contest the evidence?	Accessible review route and documented response
Who authorizes residual risk?	Named office with mandate and conflict-of-interest controls

Non-compensatory boundaries

Some conditions should not be represented as tradeable loss units. Legal rights, unacceptable safety hazards, prohibited discrimination, and mandatory procedural protections may operate as gates. If a proposed action violates a gate, the action is infeasible even if its aggregate expected loss is low. The matrix should distinguish constraints from preferences so that ethical and legal boundaries are not hidden inside a numerical score.

Participation and epistemic justice

Affected groups may possess evidence absent from administrative data and may value consequences differently from institutional decision makers. Their participation can improve the state space, loss model, and monitoring plan. Participation is not a substitute for accountable authority, but authority without relevant knowledge can produce formally elegant and substantively unjust decisions.

15 Sequential and Adaptive Decisions

Many executive decisions are not one-shot choices. They are sequences: authorize a bounded pilot, observe leading indicators, update beliefs, expand if conditions are met, and stop if harms emerge. Sequential design converts uncertainty from a reason for paralysis into a governed learning process.

Stage	Decision discipline
Authorize	Specify scope, duration, population, safeguards, and decision owner.
Observe	Collect predeclared evidence with quality controls and subgroup visibility.
Update	Recompute beliefs using the recorded evidence model; do not change definitions silently.
Review	Compare posterior expected loss, constraints, and reversal thresholds.
Act	Scale, revise, stop, wait, or escalate under the declared authority.

Pilot as an action not a delay tactic

A pilot is justified when it is feasible, reversible enough, informative, and safer than immediate scale or passive waiting. A pilot that cannot answer a decision-relevant question is merely small implementation. Predeclare the hypothesis, minimum evidence, decision threshold, stop rule, subgroup safeguards, and the action that follows each result.

Monitoring and stopping

Monitoring should target variables capable of changing the decision: posterior evidence, consequence indicators, capacity, and constraint breaches. Stop rules must be operational, not rhetorical. A valid rule specifies the metric, threshold, observation window, data owner, escalation route, and who may pause the intervention before the next scheduled review.

16 Priority Sector Applications

The mathematics transfers across sectors, but evidence generation, loss asymmetry, rights, and authority differ. The following applications are not templates to copy mechanically. They show how the same decision structure should be translated into sector-specific governance.

Sector	Uncertain state	Evidence	Asymmetric loss	Likely governed action
Agriculture	Severe crop disease	Surveillance signal and field inspection	Delayed treatment versus unnecessary ecological and fiscal cost	Targeted intervention, bounded pilot, or intensified surveillance
Education	Learners need intensive support	Assessments, attendance, teacher observation	Missed support versus labeling, opportunity cost, and inequitable allocation	Tiered support with review and subgroup audit
Health	Patient or population faces acute risk	Diagnostic test and clinical indicators	False reassurance versus unnecessary treatment and resource diversion	Treat, retest, monitor, or escalate under clinical authority
Habitat	Structure or settlement faces serious hazard	Inspection, sensor, weather, and geotechnical evidence	Delayed evacuation or retrofit versus displacement and economic disruption	Safeguard, inspect, relocate temporarily, retrofit, or escalate

Agriculture

Priors may shift rapidly with weather, pathogen mutation, seed variety, and farming practice. Evidence networks can be spatially dependent. Losses extend beyond farm revenue to food security, water quality, beneficial species, and smallholder resilience. Decision records should state whether losses are valued at farm, regional, or national level and whether affected producers can contest surveillance errors.

Education

The state is rarely a fixed learner trait. It may reflect instruction, language, attendance, disability support, household shocks, or measurement conditions. Evidence should not convert uncertainty into permanent labeling. Losses must include denied opportunity, stigma, teacher workload, and unequal access to intervention. A reversible, tiered support action is often preferable to a binary classification.

Health

Base rates and conditional probabilities are especially consequential in screening. Clinical utility depends on treatment benefit, adverse effects, delay, capacity, and patient preferences. Rights, consent, triage rules, and professional scope remain hard constraints. Population-level expected loss cannot automatically determine an individual clinical decision.

Habitat

Hazard evidence combines structural condition, exposure, vulnerability, and time. False negatives can be catastrophic; false positives can displace households and erode trust. Models must represent informal settlements and data-poor structures rather than treating missing data as low risk. Emergency authority, compensation, accessibility, and return conditions belong in the decision record.

17 Evidence Update and Decision Loss Matrix

The terminal artifact should be concise enough for authorization and complete enough for reproduction. It is a controlled record, not a decorative summary. Each field should have an owner and source.

Field	Required content
Decision and deadline	The choice, scope, population, geography, and latest responsible decision time
States	Complete-enough state set, including any severe or compound state
Actions	Feasible actions, including pilot, waiting, and escalation where relevant
Prior	Values or ranges, provenance, reference class, date, and disagreements
Evidence ledger	Observation, source, timing, quality, dependence, and likelihood model
Posterior	Calculation, model version, uncertainty, and interpretation
Loss matrix	Action-state consequences, perspective, horizon, units, and owners
Constraints	Rights, safety, legal, fiscal, capacity, and procedural gates
Bayes action	Minimum expected-loss action among feasible choices
Sensitivity	Credible ranges, reversal conditions, model alternatives, and missing-state stress test

Field	Required content
Information value	EVPI or EVSI, acquisition cost, delay loss, and net decision value
Distribution	Who benefits, who bears each error type, and contestability route
Monitoring	Indicators, thresholds, frequency, data owner, and stop rules
Authorization	ACT, PILOT, WAIT AND LEARN, or ESCALATE; named authority and review date

Authorization vocabulary

Decision	Use when
ACT	The action is feasible, robust across credible assumptions, and within authority.
PILOT	A bounded, reversible action can generate decision-relevant evidence with acceptable safeguards.
WAIT AND LEARN	Information value exceeds delay and acquisition costs, and waiting does not breach constraints.
ESCALATE	Consequences, disagreement, missing states, or constraints exceed the current authority or mandate.

18 Failure Modes

Failure	Why it fails	Correction
Ignoring the base rate	Overstates dramatic signals	Declare and test the prior reference class
Inverse-probability reasoning	Confuses likelihood with posterior	Write the full Bayes calculation
Probability-only action	Ignores asymmetric consequences	Construct the loss matrix

Failure	Why it fails	Correction
Double-counting evidence	Creates unjustified certainty	Model dependence and evidence lineage
Point-estimate theater	Hides parameter and model uncertainty	Report ranges, decision regions, and reversals
Omitted action	Forces a false binary	Include pilot, staged action, waiting, or escalation
Omitted state	Makes the posterior complete only by definition	Stress test severe, compound, and adversarial states
Aggregate-only loss	Conceals unequal burdens	Disaggregate error and delay losses
Information at any cost	Ignores delay and implementation	Calculate net information value
Model authorization	Displaces legitimate accountability	Name the human authority and constraints

Red-team questions

- Which assumption would a motivated opponent challenge first, and could that challenge reverse the action?
- What evidence source is not independent of another source already counted?
- Which group could be harmed even if the aggregate expected loss falls?
- What state would make the current action dangerously inadequate?
- What would we need to observe to stop, reverse, or escalate the decision?

19 AI Reasoning Clinic

Claim	Evaluation	Executive response
The most probable state determines the correct action.	False. Asymmetric losses can make another action optimal.	Request the loss matrix and action threshold.
The model is 90 percent accurate, so we should deploy it.	Insufficient. Accuracy can hide calibration, class imbalance, subgroup failure, and wrong losses.	Ask for decision-specific expected loss and subgroup results.
Two systems agree, so confidence should be high.	Not necessarily. They may share data, features, labels, or model	Request evidence lineage and dependence analysis.

Claim	Evaluation	Executive response
	architecture.	
More data will resolve the decision.	Only if the evidence can change action and arrives before delay costs dominate.	Request EVSI, cost, timing, and evidence-contingent actions.
A narrow credible interval means the decision is robust.	Only within the model. Missing states and misspecification can dominate parameter uncertainty.	Request alternative models and missing-state stress tests.
The algorithm recommended intervention.	Algorithms produce outputs, not legitimate authorization.	Identify the accountable human authority and governing constraints.

Using generative AI responsibly

Generative systems can help enumerate assumptions, draft alternative state spaces, translate technical material, and test the clarity of a decision record. They should not invent priors, likelihoods, loss values, or legal authority. Every quantitative input requires a source; every inference requires a reproducible method; every decision requires accountable review. Confidential or sensitive evidence must remain within authorized environments.

20 Questions Executives Should Ask Technical Teams

1. What exact decision will this analysis support, and by when?
2. What are the states and actions, and what material state or action might be missing?
3. What prior was used, whose reference class does it represent, and how old is it?
4. What does each likelihood mean, and how was it validated for this population?
5. Which evidence sources are conditionally dependent?
6. What are the false-positive, false-negative, and delay losses, and who bears each?
7. At what posterior probability does the preferred action reverse?
8. Does the recommendation survive credible alternative priors, likelihoods, and losses?
9. What are EVPI and EVSI, and are they expressed on the same loss scale as the decision?

10. What information could arrive in time to change the action?
11. Which rights, safety, legal, fiscal, or capacity constraints are non-compensatory?
12. How do results differ across relevant subgroups or locations?
13. What monitoring signal triggers continuation, revision, stopping, or escalation?
14. Who owns the model, evidence, loss specification, implementation, and final authorization?

Minimum briefing standard

A decision briefing should show the current recommendation, posterior and threshold, expected losses, key reversal assumptions, information value, constraints, distributional effects, monitoring plan, and named authority. Confidence language should describe model-conditional uncertainty and should never substitute for these elements.

21 Companion Notebook Guide

EP-M03-A08 Companion Notebook v0.2 reproduces and extends the canonical calculations. It is designed for transparent executive analysis rather than production deployment. All examples are deterministic or seeded for reproducibility.

Notebook component	Purpose	Executive interpretation
Core functions	Posterior, expected loss, Bayes action, EVPI, and EVSI	Makes the decision logic inspectable
Canonical case	Reproduces 0.36 and 0.64 posterior and 12.80 versus 2.88 losses	Confirms the lesson calculation
Sequential evidence	Updates beliefs as evidence arrives	Shows why evidence order and dependence records matter
Prior provenance	Compares credible starting beliefs	Reveals whether disagreement changes action
Likelihood uncertainty	Runs 5,000 seeded simulations	Reports how often the preferred action changes
Decision regions	Maps prior and loss combinations	Shows distance to the reversal boundary
Imperfect test	Calculates EVSI of 1.168	Tests whether further evidence is decision-relevant

Notebook component	Purpose	Executive interpretation
Missing state	Adds an emergency condition	Demonstrates when escalation is justified
Distributional view	Disaggregates consequences	Makes aggregate blind spots visible
Sector transfer	Applies the framework to four priority sectors	Separates general method from sector governance
Executive record	Exports a compact decision record	Supports review without delegating authorization

Notebook limits

The notebook is an instructional model. Its priors, likelihoods, and losses are not production estimates. Real use requires validated data, version control, access controls, independent review, sector-specific governance, and an authorized decision owner. A clean run demonstrates computational consistency, not substantive validity.

22 Practice Exercises

Exercise 1 Posterior and action

A health service estimates a 0.10 prior probability of a serious condition. A positive test has sensitivity 0.90 and false-positive probability 0.20. Calculate the posterior. Then define two actions and a defensible loss matrix. Explain why the posterior alone cannot authorize treatment.

Exercise 2 Reversal threshold

For an education intervention, the loss of no support when intensive support is needed is 15. The loss of unnecessary intensive support is 5. Derive the probability threshold above which support minimizes expected loss. Identify one rights or distributional constraint that should remain outside the compensatory loss calculation.

Exercise 3 Information value

A habitat authority can commission an inspection costing 0.8 loss units and delaying action by an expected 0.3 units. The gross EVSI is 1.4. Calculate net value and state the conditions under which the inspection should still not proceed.

Exercise 4 Robustness

Construct three coherent agriculture scenarios: historical baseline, climate-shift, and sensor-degradation. Vary the prior and likelihoods jointly. Report whether the action changes and which assumption is decision-critical.

Exercise 5 Missing state

Add a compound state to a two-state case. Explain how its inclusion affects normalization, losses, feasible actions, and authority. State whether the correct response is ACT, PILOT, WAIT AND LEARN, or ESCALATE.

Exercise 6 Build the terminal artifact

Complete an Evidence Update and Decision Loss Matrix for one decision in Agriculture, Education, Health, or Habitat. Your record must include prior provenance, an evidence ledger, posterior, loss matrix, constraints, sensitivity, information value, distributional analysis, monitoring, and named authority.

23 Knowledge Check

1. Why do base rates matter even when a signal is highly sensitive?
2. What is the difference between a likelihood and a posterior probability?
3. Why can the less probable state determine the preferred action?
4. What defines the Bayes action?
5. Why should credible alternative priors be tested rather than averaged automatically?
6. What does EVPI bound?
7. When is EVSI zero?
8. Why can two agreeing models provide less independent evidence than they appear to?
9. How does a hard constraint differ from a large loss value?
10. What conditions justify a governed pilot?
11. What is the difference between parameter uncertainty and model uncertainty?
12. Why must distributional effects be shown separately from aggregate expected loss?

Answer guide

1. They determine the starting odds; rare states can remain uncommon after a positive signal.
2. A likelihood is $P(\text{evidence} \mid \text{state})$; a posterior is $P(\text{state} \mid \text{evidence})$.
3. False-positive, false-negative, delay, and irreversible losses can be asymmetric.
4. The feasible action with minimum posterior expected loss.
5. Averaging can hide disagreement and decision reversals.
6. The maximum gross value of eliminating state uncertainty within the declared model.
7. When evidence cannot change the optimal action under the current model.
8. They may share data, assumptions, labels, or architecture.
9. A hard constraint removes an action from the feasible set; it is not traded against benefits.
10. It is bounded, reversible enough, informative, safeguarded, and linked to predeclared actions.
11. Parameter uncertainty varies values within a model; model uncertainty questions the model itself.
12. Aggregation can conceal who bears harm and whether rights or equity boundaries are breached.

24 Executive Synthesis

The governed decision chain

1. Frame the decision, deadline, population, and accountable authority.
2. Declare the state space and feasible action set, including pilot, waiting, and escalation where relevant.
3. Record the prior and its provenance; preserve credible disagreement.
4. Build the evidence ledger and likelihood model, including dependence and validation limits.
5. Update beliefs transparently and interpret the posterior as model-conditional.
6. Specify asymmetric losses, distributional effects, and non-compensatory constraints.

7. Calculate posterior expected loss and identify the Bayes action among feasible choices.
8. Test priors, likelihoods, losses, models, missing states, and reversal thresholds.
9. Compare EVPI or EVSI with acquisition cost, delay, feasibility, and rights.
10. Authorize ACT, PILOT, WAIT AND LEARN, or ESCALATE, with monitoring and review conditions.

What good looks like

A strong executive decision is not the one with the most decimals. It is the one whose assumptions are visible, whose evidence can be traced, whose losses can be contested, whose recommendation survives credible alternatives, whose information strategy is proportionate, and whose authority is legitimate. The record should make later revision easier because it shows exactly what must change before a different action becomes rational.

Terminal standard

The module is complete when the learner can defend an Evidence Update and Decision Loss Matrix before technical, operational, ethical, and authorizing audiences. The learner should be able to state not only what action is preferred, but also why, under which assumptions, how robustly, at whose risk, and with what monitoring and escalation conditions.

25 Glossary

Term	Definition
Action	A controllable choice available to the decision authority.
Bayes action	The feasible action that minimizes posterior expected loss.
Base rate	The prevalence or prior frequency of a state in a relevant reference class.
Calibration	Agreement between stated probabilities and observed frequencies across comparable cases.
Decision region	A set of input values for which the same action is preferred.

Term	Definition
EVPI	Expected value of perfect information; the upper bound on gross value from learning the state exactly.
EVSI	Expected value of sample information; the gross value of an imperfect evidence source under an optimal response policy.
Evidence ledger	A record of observations, provenance, timing, quality, dependence, and model treatment.
Expected loss	Probability-weighted average loss for an action.
Hard constraint	A non-compensatory boundary that determines feasibility.
Likelihood	The probability of observed evidence conditional on a state.
Likelihood ratio	The relative compatibility of evidence with two states.
Loss matrix	A table assigning consequence scores to action-state pairs.
Model uncertainty	Uncertainty about the structure, state space, assumptions, or applicability of a model.
Posterior	The probability distribution over states after incorporating declared evidence.
Prior	The probability distribution over states before incorporating the focal evidence.
Reversal threshold	An input boundary at which the preferred action changes.
Robust action	An action that remains acceptable across a defensible set of assumptions or models.
State	A decision-relevant condition not fully controlled or observed at the decision time.

26 References and Provenance

Core references

- Berger, J. O. Statistical Decision Theory and Bayesian Analysis. Second edition. Springer, 1985.
- Bernardo, J. M., and Smith, A. F. M. Bayesian Theory. Wiley, 1994.
- Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. Bayesian Data Analysis. Third edition. CRC Press, 2013.
- Good, I. J. Probability and the Weighing of Evidence. Charles Griffin, 1950.
- Howard, R. A. Information Value Theory. IEEE Transactions on Systems Science and Cybernetics, 1966.
- Keeney, R. L., and Raiffa, H. Decisions with Multiple Objectives. Cambridge University Press, 1993.
- Raiffa, H., and Schlaifer, R. Applied Statistical Decision Theory. Harvard University, 1961.
- Savage, L. J. The Foundations of Statistics. Wiley, 1954.

Source preservation

This canonical lesson preserves and extends the validated technical core of former Book 6 Chapter 3 Bayesian Decision Analysis and its companion notebook V6_N03_bayesian_decisions.ipynb. The preserved core includes Bayes updating, base rates, posterior beliefs, loss matrices, Bayes actions, prior and likelihood sensitivity, missing states, model misspecification, EVPI, and SRAI Principle 143.

EP-M03-A08 Companion Notebook v0.2 supplies the reproducible analytical layer used in this lesson: the agriculture calculation, sequential evidence, prior disagreement, likelihood uncertainty, decision regions, EVPI, EVSI, missing-state escalation, distributional visibility, four-sector transfer, and the executive decision record.

Looking ahead

Module 4 moves from uncertain states and asymmetric loss to comparison across plural objectives. The Evidence Update and Decision Loss Matrix produced here becomes an input to that work: it clarifies what is believed, what is at stake, which assumptions matter, and where authority or constraints limit the available alternatives.

Completion statement

The executive learner should now be able to update beliefs without overstating evidence, choose actions without confusing probability with authority, value information without ignoring delay, and preserve accountability when models remain incomplete.