Conditional Probabilities and the value of information

Outline

1. **Pricing variables contingent on events, and conditional probabilities.**
2. **A short review of a familiar result from "Bayesian" decision theory about the value of cost-free information.**
3. **Consideration of when that familiar result fails – because one of its premises is false.**

- **Moral Hazard: A simple 2-person game (from Osborne's fine textbook) where *both* players prefer that player-2 *not* learn some "cost-free" information – on the condition that, otherwise, player-1 would know of player-2's change in information.**
- **IP Dilation and the value of new information.**

1. De Finetti coherence and conditional probabilities.

De Finetti's coherence game – pricing random variables – extends to pricing a variable *contingent on an event occurring*, using this idea.

Let X be a random variable and F an event.

Aside: Recall that random variables are bounded, $\mathcal{B}$-measurable real-valued functions, and that events are the special case of 0-1 random variables.

The *Bookie* fixes a (2-sided) *contingent fair price* for X, *given* F, $q_{X|F}$.

When the *Gambler* chooses the strategy $\gamma_{X|F}$ that fixes a contract.

In state ω, the *Gambler* pays to the *Bookie* the amount

$$\gamma_{X|F}\, F(\omega)[\, X(\omega) - q_{X|F}\,].$$

So, if event F fails to occur (i.e., $F(\omega) = 0$) the contract is void. And if F occurs, the payoffs follow the usual scheme.

De Finetti established that, in order for the Bookie's 2-sided (fair) prices q_X and contingent prices, given an event F, $q_{X|F}$ to be coherent there exists a f.a. probability P where

$$q_i = \mathrm{E}_P(X_i)$$

and $$q_{X|F} = \mathrm{E}_P(X_i \mid F), \quad \text{if } P(F) > 0$$

$$q_{X|F} \quad \text{is unconstrained}, \quad \text{if } \mathrm{P}(F) = 0.$$

For indicator variables, A, B, what coherence requires is

$$\mathrm{P(AB) = P(B)P(A|B) = P(A)P(B|A)}$$

***Note well*: Contingent pricing involves no change in information.**

Neither for the *Bookie*, nor for the *Gambler*.

De Finetti's coherence criterion is a <u>static</u> constraint.

The *Bookie* offers a price for X, that yields a non-trivial contract on the condition that event F occurs.

<u>Not</u> on the condition that event G: The *Bookie* <u>*learns*</u> that F occurs.

Coherence does not entail a familiar dynamic "Bayes" learning model.

***Bayes' Rule for updating*:**

At time t_0 *YOU* have coherent, 2-sided prices and contingent prices agreeing with the probability P.

At the later time t_1 *YOU* learn (only) that event F occurs –

YOUR total new evidence is summarized by F.

Then at time t_1 YOUR updated coherent, 2-sided prices and contingent prices agree with the conditional probability $P(\mid F)$.

For ease of discussion in what follows, add to Coherence the <u>commitment</u> to use Bayes' rule for updating.

2. A basic theorem of (Bayesian) Expected Utility Theory

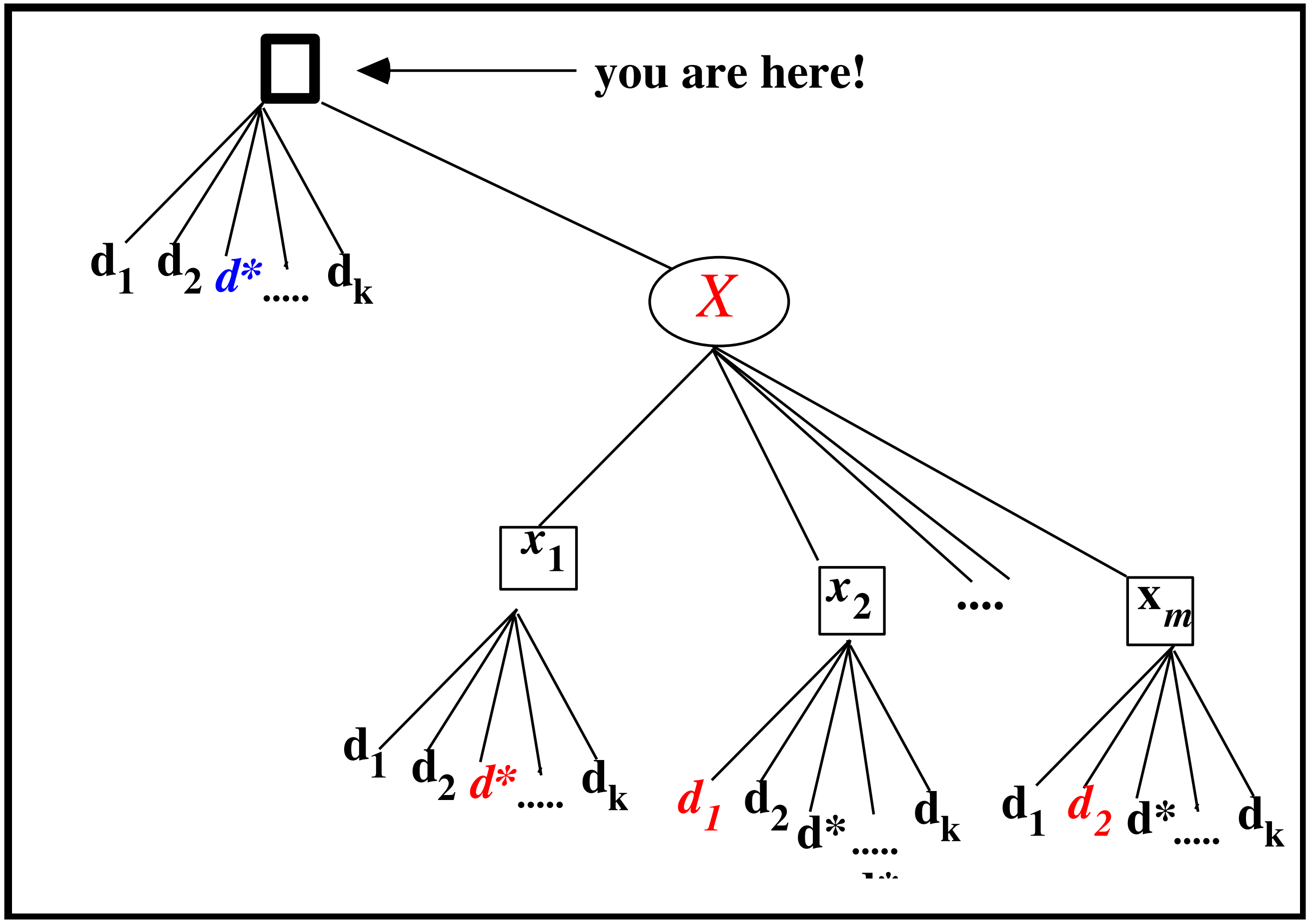
If you can postpone a terminal decision in order to observe, *cost free*, an experiment whose outcome might change your terminal decision, then it is strictly better to postpone the terminal decision in order to acquire the new evidence.

The analysis also provides a value for the new evidence, to answer:

How much are you willing to pay for the new information?

An agent faces a current decision:

- **with k terminal options $D = \{d_1, ..., d^*, ..., d_k\}$ (d^* is the best of these)**
- **and one sequential option: first conduct experiment X, with outcomes $\{x_1, ..., x_m\}$ that are observed, then choose from D.**



Terminal decisions (acts) as functions from states to outcomes

The canonical decision matrix: **decisions** × **states**

	s_1	s_2			s_j			s_n
d_1	O_{11}	O_{12}			O_{1j}			O_{1n}
d_2	O_{21}	O_{22}			O_{2j}			O_{2n}
d_k	O_{k1}	O_{k2}			O_{kj}			O_{kn}

$d_i(s_j)$ = outcome $\mathbf{o_{ij}}$.

What are "outcomes"? That depends upon which version of expected utility you consider. We will allow arbitrary outcomes, providing that they admit a von Neumann-Morgenstern cardinal utility *U*(•).

A central theme of Subjective Expected Utility [*SEU*] is this:

- **axiomatize (weak) preference $\preccurlyeq$ over decisions so that**

$$d_1 \preccurlyeq d_2 \;\; \textit{iff} \;\; \Sigma_j P(s_j)U(o_{1j}) \leq \Sigma_j P(s_j)U(o_{2j}),$$

for one subjective (personal) probability P(•) defined over *states*

and one cardinal utility U(•) defined over *outcomes*.

Then the decision rule is to choose that (an) option that *maximizes SEU*.

Note: In this version of SEU, which is the one that we will use here:

(1) decisions and states are probabilistically independent, $\mathbf{P}(s_j) = \mathbf{P}(s_j \mid d_i)$.

Reminder: This is sufficient for a general *dominance* principle.

(**2**) Utility is state-independent, $\mathbf{U_j}(o_{ij}) = \mathbf{U_h}(o_{gh})$, if $o_{ij} = o_{gh}$.

Here, $\mathbf{U_j}(o_{\bullet j})$ is the conditional utility for outcomes, given state s_j.

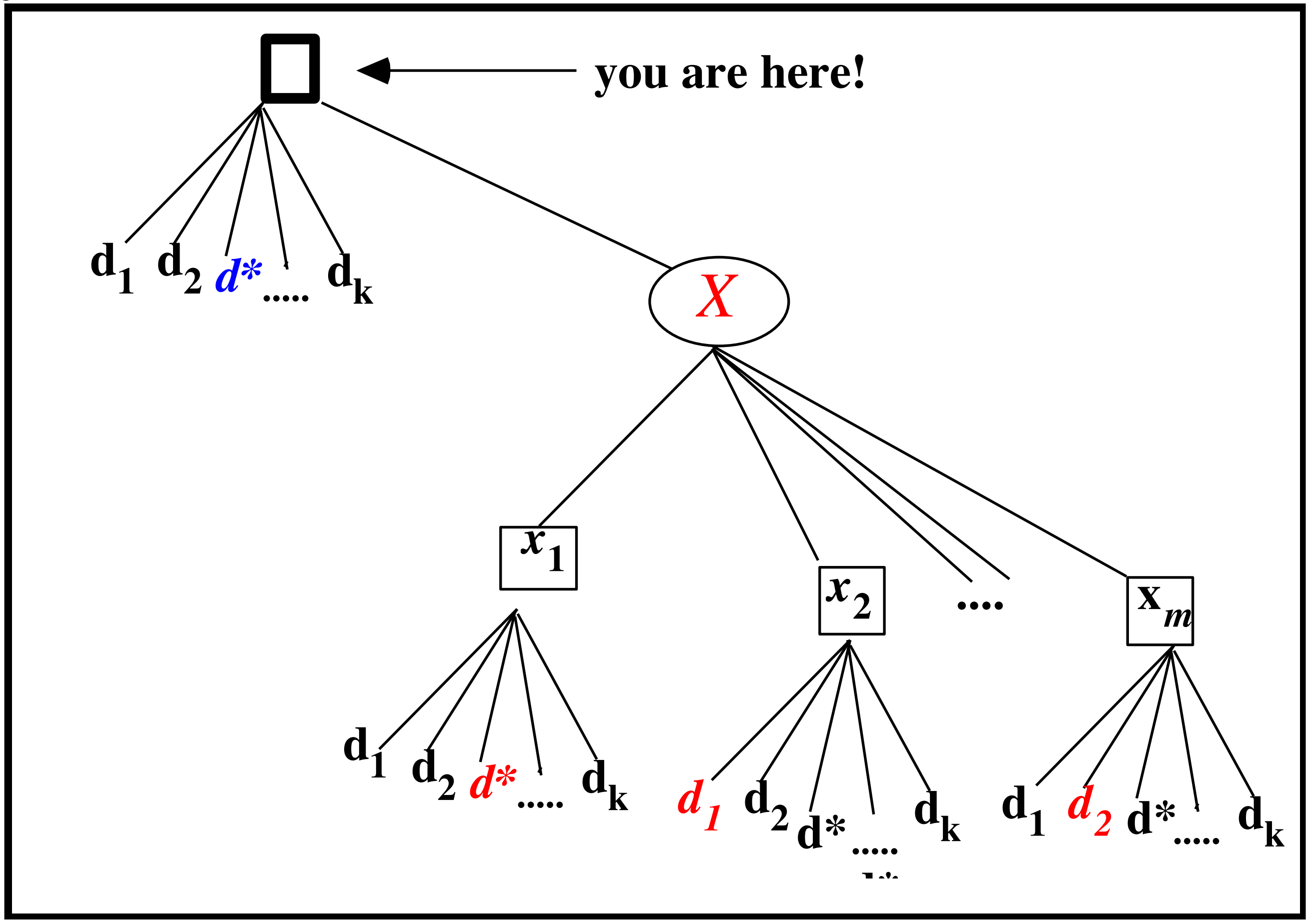
(**3**) (Cardinal) Utility is defined up to positive linear transformations,

$\mathbf{U}'(\bullet) = a\mathrm{U}(\bullet) + b$ (a > 0) is also the same utility function for purposes of *SEU*.

Note: More accurately, under these circumstances with act/state prob. independence, utility is defined up to a similarity transformation: $\mathbf{U_j}'(\bullet) = a\mathbf{U_j}(\bullet) + b_\mathbf{j}$.

So, maximizing SEU and Maximizing Subjective Expected Regret-Utility are equivalent decision rules.

Reconsider the value of cost-free evidence when decisions conform to maximizing *SEU*. Recall, the decision maker faces a choice *now* between *k*-many terminal options $D = \{d_1, ..., d^*, ..., d_k\}$ (d^* maximizes SEU among these k options). There is one sequential option: first conduct experiment *X*, with sample space $\{x_1, ..., x_m\}$, and then choose from *D* having observed *X*. Options in *red* maximize SEU at the choice nodes, using $P(s_j \mid X = x_i)$.



By the law of conditional expectations: $E(Y) = E(E[Y \mid X])$.

With Y the Utility of an option $U(d)$, and X the outcome of the experiment,

$$\begin{aligned}
\mathrm{Max}_{d \in D}\, E(U(d)) &= E(U(d^*)) \\
&= E(E(U(d^*) \mid X)) \quad \text{(“ignoring } X\text{” when choosing)} \\
&\leq E(\mathrm{Max}_{d \in D}\, E(U(d) \mid X)) \\
&= U(\text{sequential option}).
\end{aligned}$$

- **Hence, the academician's *first-principle*:**
  Never decide today what you might postpone until tomorrow in order to learn something new.

- **$U(d^*) = U(\text{sequential option})$ if and only if the new evidence X never leads you to a different terminal option.**

- **$U(\text{sequential option}) - E(U(d^*))$ is the *value of the experiment*: what you will pay (at most) in order to conduct the experiment prior to making a terminal decision.**

Example 4: Optimal Stopping: sample size, fixed versus adaptive sampling

- **See the addendum for this session's notes**

3. Failure of the value of cost-free information when there is *Moral Hazard*.

Recall: With act/state dependence even simple dominance is no longer valid!

	ω_1	ω_2
Act_1	**3**	**1**
Act_2	**4**	**2**

Regardless that Act_2 dominates Act_1, if $P(\omega_i \mid Act_i) > ¾$ then Act_1 has greater (conditional) expected utility than Act_2.

The typical model for act/state dependence is the presence of Moral Hazard, (e.g., in insurance) where the states of uncertainty for the decision maker involve the actions of another (rational) agent – as in a game!

However, regarding the principal result about the value of cost-free information, it is a side-issue whether the act/state dependence involves the actions of another decision-maker, or not.

A Toy Example of act/state dependence without Moral Hazard where new (cost free) information has negative value.

Binary Terminal Decision		ω_1	ω_2
	d_1	1	0
	d_2	0	1

Suppose $P(\omega_1) = .75$. Without added information $d^* = d_1$, and $U(d^*) = .75$.

Let $X = \{0,1\}$ be an irrelevant binary variable with likelihood,

$$P(X=0 \mid \omega_1) = P(X=0 \mid \omega_2) = .80.$$

So, X is irrelevant to Ω.

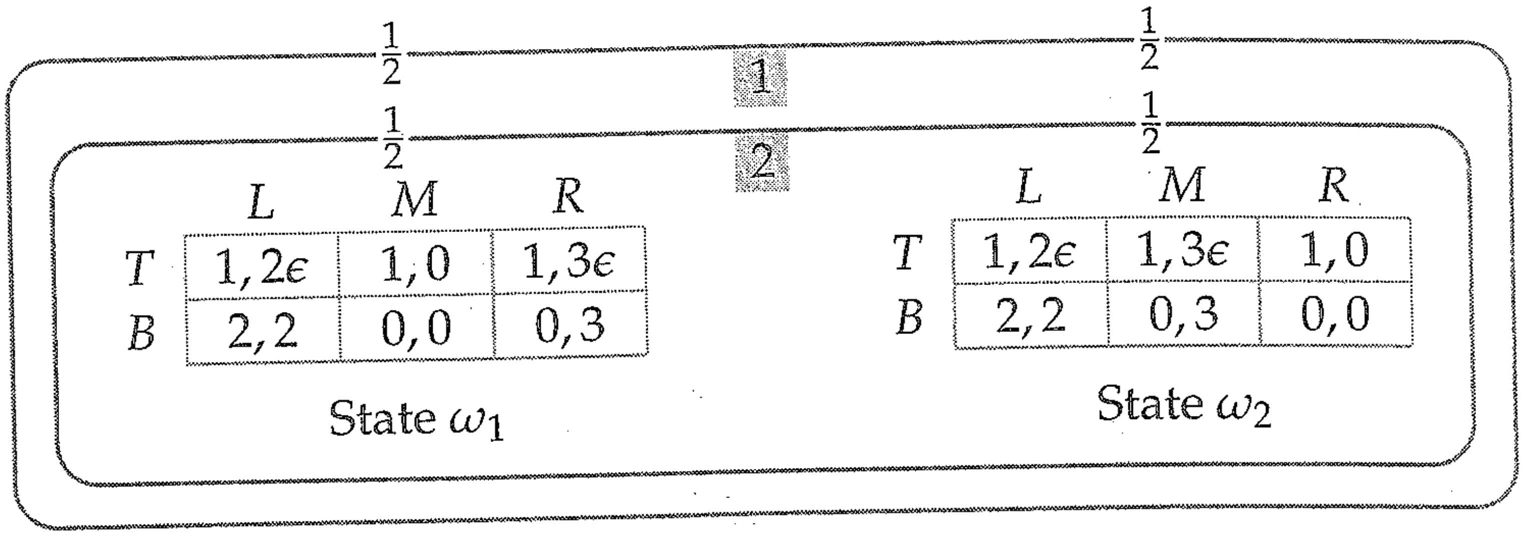
However, suppose that the decision to *observe* X alters the "prior" probability over Ω so that, $P(\omega_1 \mid \textit{observe } X) = .60 < .75$.

- **The decision to observe X creates the experiment X.**

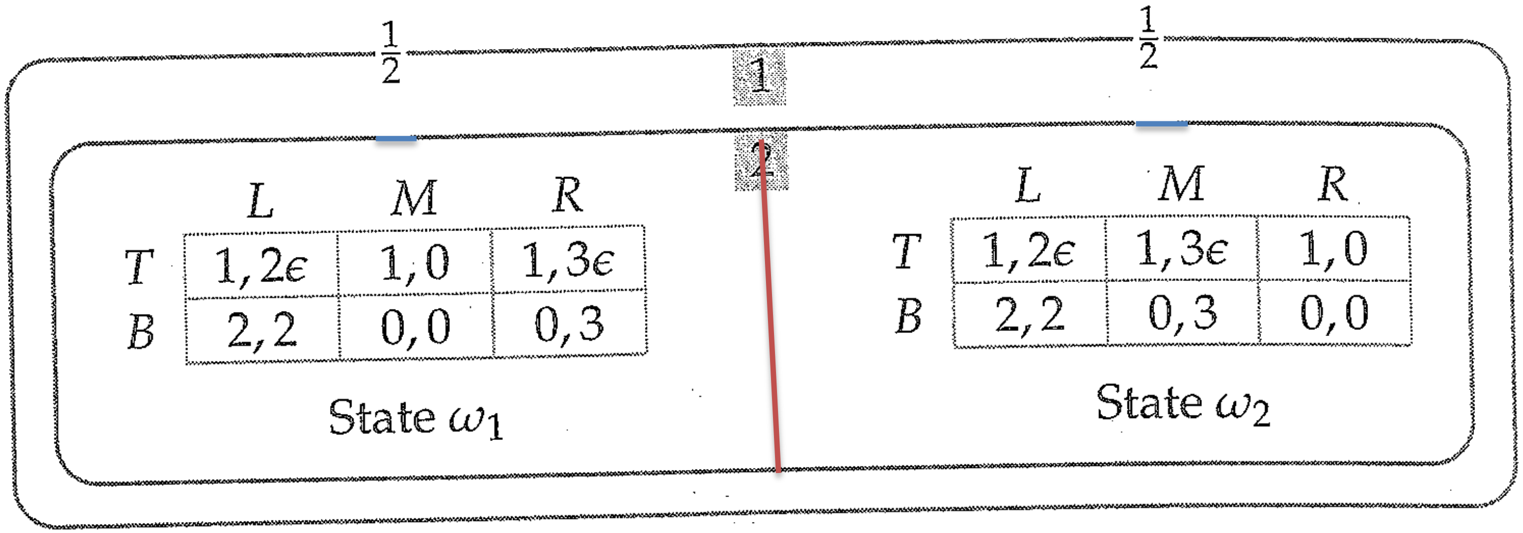
Then $U(\textit{observe } X) = .60 < 75$.

In this case, because of act/state dependence, the decision maker strictly prefers not to observe (cost free) X prior to making the terminal decision $D = \{d_1, d_2\}$.

Osborne’s game (from p. 283 of his Game Theory textbook).



This game has a unique Nash equilibrium, (B, L), with payoffs (2, 2).



	L	M	R
T	$1, 2\epsilon$	$1, 0$	$1, 3\epsilon$
B	$2, 2$	$0, 0$	$0, 3$

State ω_1

	L	M	R
T	$1, 2\epsilon$	$1, 3\epsilon$	$1, 0$
B	$2, 2$	$0, 3$	$0, 0$

State ω_2

If Column (player-2) learns the state prior to choosing, and Row knows that, then the game has a unique Nash (T, < R, L>), with payoffs (1, 3ε).

So, both players prefer the first form of the game, where Column remains ignorant of the state, and Row knows that.

	P(T & τ_1) = $\alpha/2$	P(B & τ_1) = $(1-\alpha)/2$	P(T & τ_2) = $\alpha/2$	*P(B* & τ_2) = $(1-\alpha)/2$
L	2ε	2	2ε	2
M	0	0	3ε	3
R	3ε	3	0	0

with $0 < \varepsilon \leq 1/2$

<u>Column-player's probability assumptions</u>

P(type = τ_1) = ½. P(Top) = α. P(*Row* & type) = P(*Row*)/2 – these are independent factors. Moreover, since play is simultaneous between players:

P(*Row* & type | *Column's act*) = P(*Row* & type).

Here we have act/state <u>*independence*</u> in the game with simultaneous play.

<u>Column Player's Expected Utilities for the three options</u>

U[*L*] = $2(1 - \alpha(1-\varepsilon))$ > U[*M*] = $(3/2)(1 - \alpha(1-\varepsilon))$ = U[*R*] = $(3/2)(1 - \alpha(1-\varepsilon))$

So, Column-player chooses *L*, regardless the value of α.

This is known to Row-player, who then chooses *B* to maximize her/his utility. That choice also is known to Column player; hence, <u>$\alpha = 0$</u>.

Then Column's U[*L*] = 2. Likewise, 2 is the sure payoff for Row's choice *B*.

Version 2*a* – Column-player learns her/his type prior to choosing a terminal option, and Row-player knows only that fact.

Contingent play given Column-player's type.

If Column may choose among {*L*, *M*, *R*} <u>contingent on his/her type</u>, τ_i (i =1,2)
then *R* dominates both *M* and *L*, given type = τ_1
and *M* dominates both *L* and *R*, given type = τ_2.
So the dominant contingent strategy for Column player is (*R* if τ_1, M if τ_2).
Since play is simultaneous between players, act/state independence obtains.
So the dominant play for Column has "prior" (ex ante) expected utility,

$$U[R \text{ if } \tau_1; M \text{ if } \tau_2] = 3(1 - \alpha'(1-\varepsilon)),$$

where α' is Column player's "prior" for Row choosing *Top* in Version 2a.

In Version 1, Column's P(Top) = α. If $\alpha = \alpha'$, then

$$U[R \text{ if } \tau_1; M \text{ if } \tau_2] = 3(1 - \alpha(1-\varepsilon)) > U[L] = 2(1 - \alpha(1-\varepsilon))$$

and Column player has <u>positive value</u> for the information of her/his type, all in accord with the Basic Result.

HOWEVER, in version 2a of the problem, since Row knows these calculations on behalf of Column, and as Row's option *T* dominates option *B* given either *M* or R – with payoffs 1 *vs* 0 -- then, Row chooses *T*, and Column knows this too.

- **So, <u>$\alpha' = 1$</u> $\neq$ $\alpha = 0$ and we have act/state dependence (for Column)**

In Version 2*b* of the game both players learn Column's type prior to making a terminal decision. The upshot is the same.

By dominance, Column plays: *R* if τ_1; *M* if τ_2. Knowing this Row plays *T*, etc.

From Column's perspective, in Version 2*a* (or 2*b*),

U[*R* if τ_1; *M* if τ_2)] = 3ε < 2 = Version 1's U[*L*].

Column prefers the first version of the game. Similarly for Row player!

So, if the <u>*initial choice*</u> (for either player to make) is whether to play Version 1, or instead to play Version 2a (or 2b) of the game, the initial choice is to play Version 1 of the game.

In this sequential problem, in choosing first between Version 1 and Version 2 of the game, and then playing the version chosen, there is act/state dependence from either player's perspective: probabilistic dependence between the player's choice of Version 1 vs. Version 2 of the game and her/his probability for how the other player chooses.

From Column-player's perspective, the mere choice of version fixes the value of α – Column player's probability that Row player chooses Top, *T*.

Likewise, in choosing between Version 1 and Version 2 of the game, Row player faces act/state dependence in her/his probability for Column's behavior.

Thus, the familiar result about the non-negative value of cost-free information does not apply in this sequential game.

Each player prefers Version 1 over Version 2.
Each player prefers playing the game with less information rather than more.
And that is explained by the presence of act/state dependence – for each player.

But this same phenomenon can happen when there is only one decision maker and she/he faces a problem with act/state dependence in probabilities.
The opportunity to postpone a cost-free decision may have negative value (with or without the Moral Hazard of *<u>another</u>* decision maker's choice) provided that there is act/state dependence in personal probabilities.

Note well: Though the result about the non-negative value of cost-free information is not robust over situations with act/state dependence, nonetheless:

- **de Finetti's coherence is robust over situations with act/state dependence.**

	ω_1	ω_2	…	ω_k	…
Act_1	o_{11}	o_{12}	…	o_{1k}	…
Act_2	o_{21}	o_{22}	…	o_{2k}	…

Defn.: **Say that Act_2 *robustly dominates* Act_1 if Act_2 *uniformly dominates* Act_1 and *supremum* $\{U(o_{1j})\}$ < *infimum* $\{U(o_{2j})\}$.**

Then regardless the act/state dependence, i.e. regardless $P_1(\omega)$ and $P_2(\omega)$,

$$\sum_j P_1(\omega_j)U(o_{1j}) < \sum_j P_2(\omega_j)U(o_{2j})$$

and **Act_1 is strictly dispreferred to Act_2.**

When the *Bookie* has an incoherent set of fair-prices, and when the *Gambler* uses the strategy $\{\gamma^*_i\}$, the *Bookie* suffer a uniform sure-loss ($o_{ij} < e^* < 0$) compared with *Abstaining*.

	ω_1	ω_2	…	ω_k	…
Incoherent pricing	o_{11}	o_{12}	…	o_{1k}	…
Abstain from playing	*0*	*0*	…	*0*	…

- **Abstaining robustly dominates incoherent pricing.**

There is no salvation for the Bookie from such incoherence even by allowing collusion between *Nature* (which determines ω) and the *Bookie*!

4. Negative Value of Cost-Free information within IP theory.

The decision rule is not Expected Utility Maximization with a single probability distribution. For example, represent uncertainty of an event using a (convex) <u>set of probabilities</u>, $\mathcal{P}$. Let the decision rule be

Γ-Maximin – choose an act whose min expected utility is max w.r.t. set $\mathcal{P}$.

Then the value of (cost free) information may be negative.

- **This is the fate of inference with *pivotal variables* in statistical inference.**

Next, I illustrate this situation when there is *Dilation* for sets of probabilities.

Dilation for sets of probabilities.

Let $\mathcal{P}$ be a (convex) set of probabilities on algebra $\mathcal{A}$.

For an event E, denote by

$P_*(\mathrm{E})$ the _lower_ probability of E: $inf_{\mathrm{P}\in\mathcal{P}}$ {P(E)}

and $P^*(\mathrm{E})$ the _upper_ probability of E: $sup_{\mathrm{P}\in\mathcal{P}}$ {P(E)}.

Let $X = (x_1, ..., x_n)$ be a partition, here taken to be finite for simplicity.

The set of conditional probabilities {P(E | x_i)} (strictly) _dilate_ if

$$P_*(\mathrm{E} \mid x_i) < P_*(\mathrm{E}) \leq P^*(\mathrm{E}) < P^*(\mathrm{E} \mid x_i)$$

for each $i = 1, ..., n$.

That is, dilation occurs provided that, for each event ($X = x_i$) in a partition,

the set of conditional probabilities for an event E, given x_i, properly include the unconditional probabilities for E.

Dilation of conditional probabilities is the opposite phenomenon to the more familiar "shrinking" of sets of opinions with increasing shared evidence.

Example **– also illustrating when normal and extensive form decisions differ**

Suppose A is a highly *IP-uncertain* event. That is $P^*(A) - P_*(A) \approx 1$.

Let {H,T} indicate the flip of a fair coin whose outcomes are independent of A.

That is, $$P(A,H) = P(A)/2 \text{ for each } P \in \mathcal{P}.$$

Define event E by, $E = \{(A,H), (A^c,T)\}$.

Evidently, $P(E) = .5$ for each $P \in P$.

Note: E is a "pivotal" variable involving A and the coin flip.

	H	T
A	E	E^c
A^c	E^c	E

Then $$0 \approx P_*(E \mid H) < P_*(E) = .5 = P^*(E) < P^*(E \mid H) \approx 1$$

and $$0 \approx P_*(E \mid T) < P_*(E) = .5 = P^*(E) < P^*(E \mid T) \approx 1.$$

Thus, regardless how the coin lands, conditional probability for event E dilates to a large interval, from a determinate value .5.

This example mimics Ellsberg's (1961) *paradox*, where the mixture of two *uncertain* events has a determinate probability.

Consider a sequential (extensive form) choice between:

Terminal option d_1 — Win \$.75 if E and Lose \$1.25 if E^c,

and a Sequential option X - observe the coin flip {H, T} and then choose

between d_2 — an even money \$1 bet on E. (Note: $d_1 = d_2 -$ \$.25 *fee*.)

and d_3 — a *fee* of \$.50.

In the normal form of this problem, there are 5 options (including four combinations of how to choose between d_2 and d_3 given X).

In a pairwise choice between d_1 and d_2, d_2 (simply) dominates option d_1.
In the normal form, each of d_1 and d_2 have determinate expected values

- **The Expected Utility(d_1) = -1/8 and the Expected Utility(d_2) = 0.**

In the normal form d_1 is not *E*-admissible (nor is it Γ-maximin admissible) as d_1 fails to maximize expected utility for each $P \in \mathcal{P}$.

Each of the 3 normal-form options involving d_3 likewise is inadmissible against d_2.

Option d_2 is the unique admissible (normal form) option, with value 0.

In the sequential (extensive form) problem, conditional upon either H or T,

both choices d_2 and d_3 are (pairwise) E-admissible.

However, given H or T, d_3 maximizes minimum expected utility in a pairwise choice with d_2.

Given H (or T), d_3 has a higher, minimum conditional expected (Γ-maximin) value, equal to - ½, than does d_2.

The minimum conditional expected (Γ-maximin) value of d_2 equals -1.

So, under the sequential option X, d_3 alone is chosen.

Then, by backward induction, a contrast at the initial node between d_1 and d_3 reveals that d_1 maximizes expected utility, with value -1/8.

- **The extensive form Γ-maximin admissible option is inadmissible in the normal form.**
- **In this problem, the observation X has a negative value of -1/8 !**

Dilation and Independence.

Independence is sufficient for dilation.

Let Q be a convex set of probabilities on algebra A and suppose we have access to a fair coin which may be flipped repeatedly: algebra C for the coin flips.

Assume the coin flips are mutually independent and, with respect to Q, also independent of events in A.
Let P be the resulting convex set of probabilities on $A \times C$

(This condition is similar to, e.g., DeGroot's assumption of an extraneous continuous r.v., and is similar to the "fineness" assumptions in the theories of Savage, Ramsey, Jeffrey, etc.)

Theorem: If Q is not a singleton, there is a 2×2 table of the form $(E,E^c) \times (H,T)$ where both:

$$P_*(E \mid H) < P_*(E) = .5 = P^*(E) < P^*(E \mid H)$$
$$P_*(E \mid T) < P_*(E) = .5 = P^*(E) < P^*(E \mid T).$$

That is, dilation occurs.

Independence is necessary for dilation.

Let P be a convex set of probabilities on algebra A. The next result is formulated for subalgebras of 4 atoms: (p_1, p_2, p_3, p_4)

The case of 2×2 tables.

	B_1	B_2
A_1	p_1	p_2
A_2	p_3	p_4

Define the quantity
$\mathbf{S}_{\mathrm{P}}(A_1,B_1) = p_1/(p_1+p_2)(p_1+p_3) = P(A_1,B_1) / P(A_1)P(B_1)$.

Thus, $\mathbf{S}_{\mathrm{P}}(A_1,B_1) = 1$ iff A and B are independent under P.

Lemma: If P displays dilation in this sub-algebra, then
$$\inf_{\mathrm{P}}\{\mathbf{S}_{\mathrm{P}}(A_1,B_1)\} < 1 < \sup_{\mathrm{P}}\{\mathbf{S}_{\mathrm{P}}(A_1,B_1)\}.$$

Theorem: If P displays dilation in this sub-algebra, then there exists $P^{\#} \in$ P such that $\mathbf{S}_{\mathrm{P}^{\#}}(A_1,B_1) = 1$.

Dilation and the ε-contaminated model.

Given probability P and $1 > \varepsilon > 0$, define the convex set

$\mathcal{P}_{\varepsilon}(P) = \{(1-\varepsilon)P + \varepsilon Q$: Q an arbitrary probability$\}$.

This model is popular in studies of Bayesian Robustness.

(See Huber, 1973, 1981; Berger, 1984.)

Also, it is equivalent to the model formed by fixing effective lower probabilities for the atoms of an algebra.

Lemma In the ε-contaminated model, dilation occurs in an algebra $\mathcal{A}$ iff it occurs in some 2×2 subalgebra of $\mathcal{A}$.

So, the next result is formulated for 2x2 tables.

$\mathcal{P}_\varepsilon(P)$ experiences dilation *if and only if*

case 1: $S_P(A_1,B_1) > 1$

$$\varepsilon > [S_P(A_1,B_1) - 1] \times \max\{ P(A_1)/P(A_2) ;\ P(B_1)/P(B_2) \}$$

case 2: $S_P(A_1,B_1) < 1$

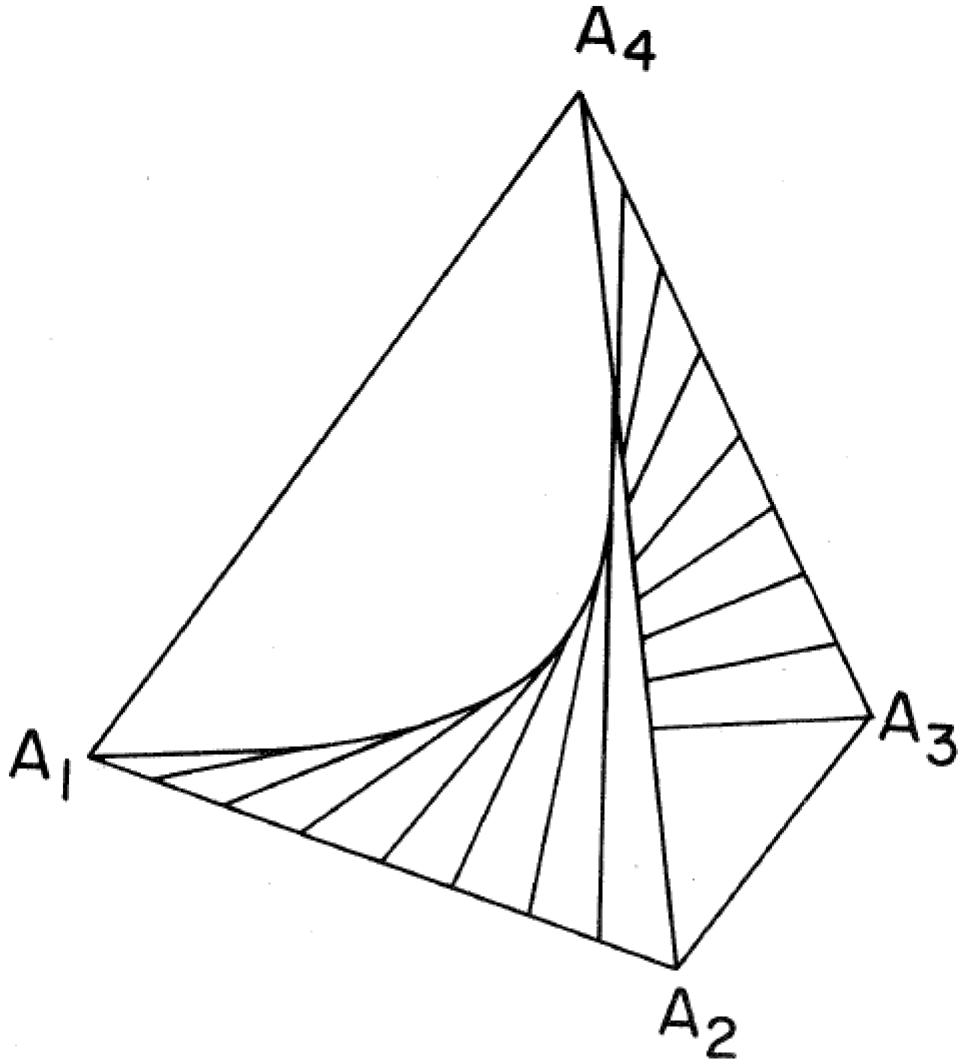
$$\varepsilon > [1 - S_P(A_1,B_1)] \times \max\{ 1 ;\ P(A_1)P(B_1)/P(A_2)P(B_2) \}$$

and case 3: $S_P(A_1,B_1) = 1$

P is internal to the simplex of all distributions.

Thus, dilation occurs in the ε-contaminated model if and only if the focal distribution, P, is close enough (in the tetrahedron of distributions on four atoms) to the saddle-shaped surface of distributions which make A and B independent.

Here, S_P provides one relevant index of the proximity of the focal distribution P to the surface of independence.



	F	F^c
E	A_1	A_2
E^c	A_3	A_4

SUMMARY for Session 2.

We reviewed a basic result about the value of new information, when cost-free information has non-negative value relative to a terminal decision.

We considered two departures from this basic result:

(1) When the agent's opinion has act/state probabilistic dependence, then cost-free evidence may have negative value. We saw how this may arise in a simple 2-person Bayesian game – Osborne's game.

(2) When the agent uses, e.g., IP-decision making – then *dilation* results in cost-free evidence wit

Addendum: Experiment Design and Optimal Stopping

***Example 4*: Choosing sample size, fixed versus adaptive sampling**

See the addendum for this session's notes

(DeGroot, chpt. 12)

The statistical problem has a terminal choice between two options, $D = \{ d_1, d_2\}$.

There are two states $S = \{s_1, s_2\}$, with outcomes that form a *regret* matrix:

$U(d_1(s_1)) = U(d_2(s_2)) = 0$, $U(d_1(s_2)) = U(d_2(s_1)) = -b < 0$.

	s_1	s_2
d_1	0	$-b$
d_2	$-b$	0

Obviously, according to SEU, $d^* = d_i$ if and only if $P(s_i) \geq .5$ $(I = 1, 2)$.

Assume, for simplicity that $P(s_1) = p < .5$, so that $d^* = d_2$ with $E(U(d_2)) = -pb$.

***The sequential option*: There is the possibility of observing a random variable $X = \{1, 2, 3\}$. The statistical model for X is given by:**

$$P(X = 1 \mid s_1) = P(X = 2 \mid s_2) = 1 - \alpha.$$

$$P(X = 1 \mid s_2) = P(X = 2 \mid s_1) = 0.$$

$$P(X = 3 \mid s_1) = P(X = 3 \mid s_2) = \alpha.$$

Thus, $X = 1$ or $X = 2$ identifies the state, which experimental outcome has conditional probability 1-α on a given trial; whereas $X = 3$ is an irrelevant datum, which occurs with (unconditional) probability α.

Assume that X may be observed repeatedly, at a cost of c-units per observation, where repeated observations are conditionally *iid*, given the state s.

- ***First*, we determine what is the optimal fixed sample-size design, $N = n^*$.**
- ***Second*, we show that a sequential (adaptive) design is better than the best fixed sample design, by limiting ourselves to samples no larger than n^*.**
- ***Third*, we solve for the global, optimal sequential design as follows:**
  - **We use Bellman's principle ("backward induction) to determine the optimal sequential design bounded by $N \leq k$ trials.**
  - **By letting $k \to \infty$, we solve for the global optimal sequential design in this decision problem.**

- ***The best, fixed sample design*.**

Assume that we have taken $n > 0$ observations: $\tilde{X} = (x_1, \ldots, x_n)$

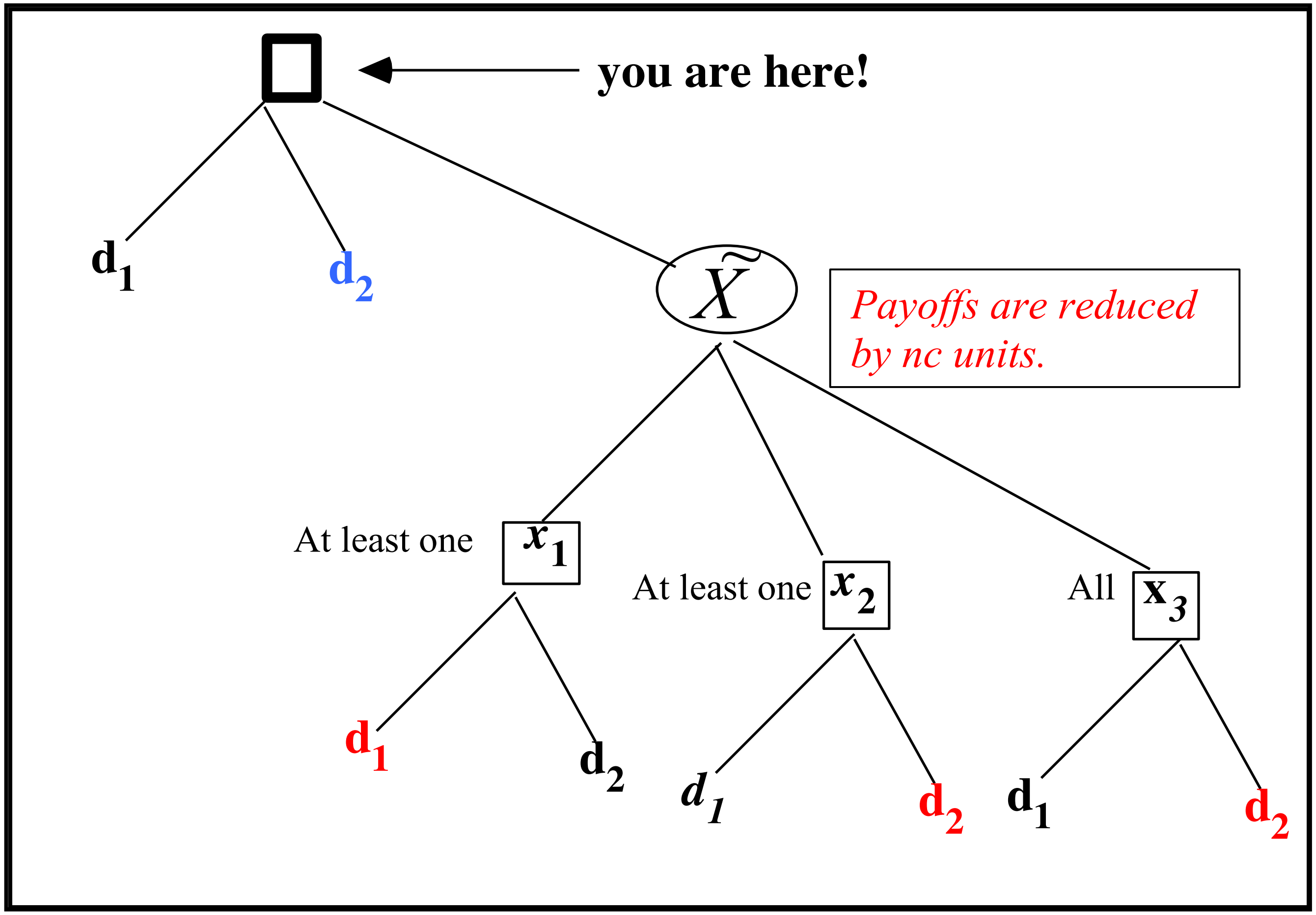
The posterior prob., $P(s_1 \mid \tilde{X}) = 1$ ($P(s_2 \mid \tilde{X}) = 1$ $x_i = 2$) if $x_i = 1$ for some $I = 1, \ldots, n$. Then, the terminal choice is made at no loss, but nc units are paid out for the experimental observation costs.

Otherwise, $P(s_1 \mid \tilde{X}) = P(s_1) = p$, when all the $x_i = 3$ ($I = 1, \ldots, n$), which occurs with probability α^n. Then, the terminal choice is the same as would be made with no observations, d_2, having the same expected loss, $-pb$, but with nc units paid out for the experimental observation costs.

That is, the pre-trial (SEU) value of the sequential option to sample n-times and then make a terminal decision is:

$$E(\textbf{sample } n \textbf{ times before deciding}) = -[pb\alpha^n + cn].$$

Assume that c is sufficiently small (relative to $(1-\alpha)$, p and b) to make it worth sampling at least once, i.e. $-pb < -[pb\alpha + c]$**, or** $c < (1-\alpha)pb$



Thus, with the pre-trial value of the sequential option to sample *n*-times and then make a terminal decision:

$$E(\textbf{sample } n \textbf{ times before deciding}) = -[pb\alpha^n + cn].$$

- **then the *optimal fixed sample size design* is, approximately (obtained by treating *n* as a continuous quantity):**

$$n^* = \frac{-\log[pb\log(1/\alpha)/c]}{1/\log(1/\alpha)}$$

- **and the *SEU* of the optimal fixed-sample design is approximately**

$$E(\textbf{sample } n^* \textbf{ times then decide}) = -(c/\log(1/\alpha))\,[1 + \log[pb\log(1/\alpha)/c]\,]$$

$$> -pb = E(\textbf{decide without experimenting})$$

- **Next, consider the plan for bounded sequential stopping, where we have the option to stop the experiment after each trial, up to n^* many trials.**

At each stage, n, prior to the $n^{*\text{th}}$, evidently, it matters for stopping only whether or not we have already observed $X = 1$ or $X = 2$.

- **For if we have then we surely stop: there is no value in future observations.**
- **If we have not, then it pays to take at least one more observation, if we may (if $n < n^*$), since we have assumed that** $c < (1-\alpha)pb$.

If we stop after n-trials ($n < n^*$), having seen $X = 1$, or $X = 2$, our loss is solely the cost of the observations taken, nc, as the terminal decision incurs no loss.

Then, the expected number of observations N from bounded sequential stopping (which follows a *truncated negative binomial* distn) is:

$$E(N) = (1-\alpha^{n^*})/(1-\alpha) \; < \; n^*.$$

Thus, the *S*ubjective *E*xpected *U*tility of (bounded) sequential stopping is:

$$-[pb\alpha^{n^*} + cE(N)] \; > \; -[pb\alpha^{n^*} + cn^*].$$

- **What of the unconstrained sequential stopping problem?**

With the terminal decision problem $D = \{ d_1, d_2\}$, what is the global, optimal experimental design for observing X subject to the constant cost, c-units/trial and the assumption that $c < (1-\alpha)pb$?

Using the analysis of the previous case, we see that if the sequential decision is for bounded, optimal stopping, with $N \leq k$, the optimal stopping rule is to continue sampling until either $X_i \neq 3$, or $N = k$, which happens first. Then, we see that

$E_{N\leq k}(N) = (1-\alpha^k)/(1-\alpha)$ and the SEU of this stopping rule is $-[pb\alpha^k + c(1-\alpha^k)/(1-\alpha)]$,

which is monotone increasing in k.

Thus the global, optimal stopping rule is the unbounded rule: continue with experimentation until $X = 1$ or $= 2$, which happens with probability 1.

$E(N) = 1/(1-\alpha)$ and the SEU of this stopping rule is $-[c/(1-\alpha)]$.

Note: Actual costs here are unbounded!

The previous example illustrates a basic technique for finding a global optimal sequential decision rule:

1) Find the optimal, *bounded* decision rule $\boldsymbol{d}_{\boldsymbol{k}}^{*}$ when stopping is mandatory at $\boldsymbol{N} = \text{k}$.

In principle, this can be achieved by *backward induction*, by considering what is an optimal terminal choice at each point when $N = \text{k}$, and then using that result to determine whether or not to continue from each point at $N = \text{k-1}$, etc.

2) Determine whether the sequence of optimal, bounded decision rules converge as $\text{k}\to\infty$, to the rule $\boldsymbol{d}_{\infty}^{*}$.

3) Verify that $\boldsymbol{d}_{\infty}^{*}$ is a global optimum.