The $64 Question

- Consider the problem faced by a contestant on a TV quiz show, in which there are six stages (1, …6).
- At any stage, the person may choose to quit and receive payoff $\$2^{i-1}$, with a top prize of $64.
- If the person chooses to continue, she is presented with a question which, if correctly answered, allows her to advance to the next stage (i+1), but if not correctly answered, forces her to quit with no payoff.
- The questions become progressively more difficult at each stage, of course, and she estimates that the probability that she can answer the question at stage i to be P_i where $P_{i+1}<P_i$.

1. Formulate a dynamic programming model to compute her optimal strategy.

   - What are the states? ______
   - What is the decision set for each state? ________
   - What is the recursive definition of the optimal value function?

State Vector

i	s[i]	name
1	1	Active
2	0	Stopped

Decision Vector

i	x[i]	name
1	1	Continue
2	0	Stop

Random Variable

i	d[i]	name
1	1	Success
2	0	Failure

Optimal Value Function:

$f_n(s)$ = maximum expected reward if at stage n the current state is s

Recursive definition:

$$f_n(0) = f_{n+1}(0) \quad \forall n = 1, 2, \ldots 6$$

$$f_n(1) = \max \begin{cases} \mathrm{R}[\mathrm{n}\text{-}1] & \sim \ \mathrm{x}=0 \text{ (stop)} \\ \mathrm{p_n} f_n(1) + (1 - p_n) f_n(0) & \sim \ \mathrm{x}=1 \text{ (continue)} \end{cases}$$

APL implementation of Optimal Value Function

```
     ’ z„F N;t
[1]    ©
[2]    © Optimal Value Function
[3]    ©   for optimal stopping problem
[4]    ©
[5]    :if N>NN
[6]      z„(Reward[NN+1]),0,-BIG
[7]    :else
[8]      © Recursive definition of optimal value function
[9]      z„(P[N],1-P[N]) Maximize_E
   ((s×Reward[N])°.×(1-x)°.+0×d)+(F N+1)[TRANSITION s°.×x°.×d]
[10]   :endif
     ’
```

2. Specify values for P_i, i=0,1, …6 and compute the optimal strategy.

Stage	***1***	***2***	***3***	***4***	***5***	***6***
P{success}	0.8	0.7	0.6	0.5	0.4	0.3

```
Reward „  1 2 4 8 16 32 64
```

```
                                 Quiz Show

Recursion type: forward

                    ---Stage 6---
    s  \ x:    1         0   | Maximum
    1  |   19.2000   32.0000|   32.0000
    0  |    0.0000    0.0000|    0.0000
```

Final Stage

Example calculation:

If s=1 and x=1, i.e., the contestant is still active and chooses to continue, the expected reward is

$$0.3\times f_7(1)+0.7\times f_7(0)$$

$$0.3\times 64+0.7\times 0 \quad = \quad 19.20$$

```
---Stage 5---
     s   \ x:     1          0   | Maximum
----------------------------------------------
     1   |   12.8000   16.0000|   16.0000
     0   |    0.0000    0.0000|    0.0000

---Stage 4---
     s   \ x:     1          0   | Maximum
----------------------------------------------
     1   |    8.0000    8.0000|    8.0000
     0   |    0.0000    0.0000|    0.0000

---Stage 3---
     s   \ x:     1          0   | Maximum
----------------------------------------------
     1   |    4.8000    4.0000|    4.8000
     0   |    0.0000    0.0000|    0.0000

---Stage 2---
     s   \ x:     1          0   | Maximum
----------------------------------------------
     1   |    3.3600    2.0000|    3.3600
     0   |    0.0000    0.0000|    0.0000

---Stage 1---
     s   \ x:     1          0   | Maximum
----------------------------------------------
     1   |    2.6880    1.0000|    2.6880
     0   |    0.0000    0.0000|    0.0000
```

Summary of Optimal Returns and Decisions

Stage 6

Current State	Optimal Decision	Optimal Value
Active	Stop	32.0000
Stopped	Continue Stop	0.0000

Stage 5

Current State	Optimal Decision	Optimal Value
Active	Stop	16.0000
Stopped	Continue Stop	0.0000

Stage 4

Current State	Optimal Decision	Optimal Value
Active	Continue Stop	8.0000
Stopped	Continue Stop	0.0000

Stage 3

Current State	Optimal Decision	Optimal Value
Active	Continue	4.8000
Stopped	Continue Stop	0.0000

Stage 2

Current State	Optimal Decision	Optimal Value
Active	Continue	3.3600
Stopped	Continue Stop	0.0000

Stage 1

Current State	Optimal Decision	Optimal Value
Active	Continue	2.6880
Stopped	Continue Stop	0.0000

The **optimal strategy** is therefore to continue playing until the contestant fails a question or reaches stage 4, at (assuming she is risk-neutral) she is indifferent toward stopping or continuing.

At stage 5, if she is still active, she should quit.

Expected reward at beginning of game: **$2.688**