The grading problem and optimal stopping

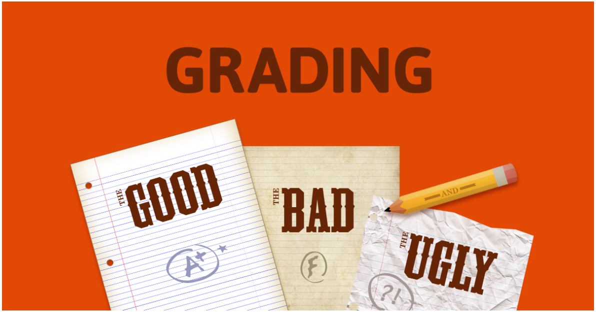
Yuqiong Wang

February 25, 2021

Outline

Problem Setting & The Bayesian approach

Problem Setting



1. Suppose there are two types of students: good, and bad.
2. Each student needs to solve 40 problems.
3. For each correct solution, they get $+1$ point, for each wrong solution they get 0 point.
4. Our goal is to decide if we pass or fail a student.

The Bayesian approach

- Are all the students who got under 18 points bad students?
- Is there any information we are not using?

Bayes' theorem

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B|A)\mathbb{P}(A) + \mathbb{P}(B|A^C)\mathbb{P}(A^C)}.$$

Back to our problem

What we know about the group

- Suppose there are 50% of good students,
- a good student solves a problem correctly with probability p,
- a bad student solves a problem correctly with probability $q, q < p$.

We define for each student:

- Y_i: the score on the ith problem, $Y_i \in \{0, 1\}$,
- X: the total score, $X = \sum_{i=1}^{40} Y_i$,
- d: the decision, $d \in \{G, B\}$,
- Π_G^X: the probability that the student is good,
- Π_B^X: the probability that the student is bad.

Back to our problem

We want to make a decision (G/B) to minimise the probability of making a mistake.

$$V_N = \inf_d \{\mathbb{P}(d = G, B) + \mathbb{P}(d = B, G)\}.$$

Now what is the rational decision after knowing X?

How does the value function change with the prior belief?

Assume an arbitrary student is "good" with probability π.

Goal: Find d^*, such that V_N is minimised.

In general, the value function can be regarded as a function of the prior π.

Optimise over deterministic times

What if grading is not free: new formulation

- Assume each student answers infinitely many questions,
- We can choose the number of problems we grade
- Grading is not free, it costs c per problem.
- At time 0, we decide when to stop.

Our value function becomes

$$\bar{V} = \inf_{d,T} \{\mathbb{P}(d = G, B) + \mathbb{P}(d = B, G) + cT\}.$$

Goal: Find a pair (d^*, T^*), such that $\bar{V}$ is minimised.

Step 1: Value function

Similarly, the value function can be written as

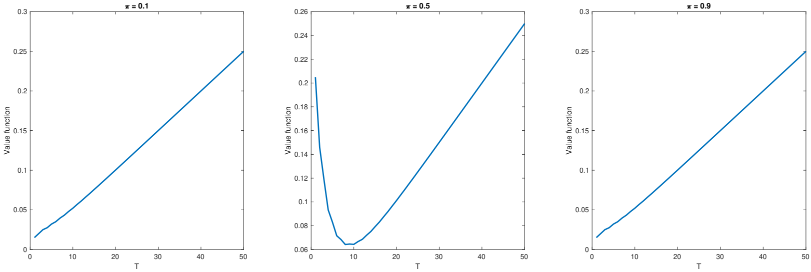
$$\bar{V}(\pi) = \inf_{T}\{\mathbb{E}_{\pi}[\Pi_T \wedge (1 - \Pi_T) + cT]\},$$

where Π_T is a **stochastic process** with $\Pi_0 = \pi$. For each $t \in \{1, 2, \ldots, T\}$,

Observation: Π_T is a **Markov process**.

Step 2: find the strategy T^*

With $p = 0.8, q = 0.2, c = 0.005$, plot $\mathbb{E}_\pi[\Pi_T \wedge (1 - \Pi_T) + cT]$ w.r.t. T:



- How does the prior belief affect T^*?
- What do you do if c is large?

Optimise over stopping times

Can we do better?

Yes! By optimising over **stopping times**.

Stopping times

Let $\mathcal{F}_t^X$ be the set of all the information generated by process X up to time t, τ is a stopping time if

$$\{\tau \leq t\} \in \mathcal{F}_t^X.$$

- τ is a random variable,
- τ cannot depend on the future,
- Every deterministic t is a stopping time.

(What does the last bullet point tell us?)

Formulating the problem (again)

Denote $g(\pi) = \pi \wedge (1 - \pi)$, then

$$V(\pi) = \inf_{\tau} \mathbb{E}_{\pi}[g(\Pi_{\tau}) + c\tau],$$

where Π is the same stochastic process as before.

<u>Observation 1:</u>

$$V(\pi) \leq \bar{V}(\pi).$$

<u>Observation 2:</u>

$$V(\pi) \leq g(\pi).$$

General theory

Optimal stopping

In optimal stopping problems, we consider sequentially observed random variables and determine the optimal time for taking a certain action, in order to maximise the expected gain (or minimise the expected cost).

$$V(x) = \sup_{\tau} \mathbb{E}_x[G(X_\tau)]$$

Observations:

- Need to find the value function in order to find its associated τ^*
- It is always better than optimising over deterministic times
- It is always better than stopping now.

Step 1: Find the value function

Idea: Standing at each time point, you have to options:

- Stop now,
- Continue for one more step, pay the cost, and face the same choice again.

Define an operator T:

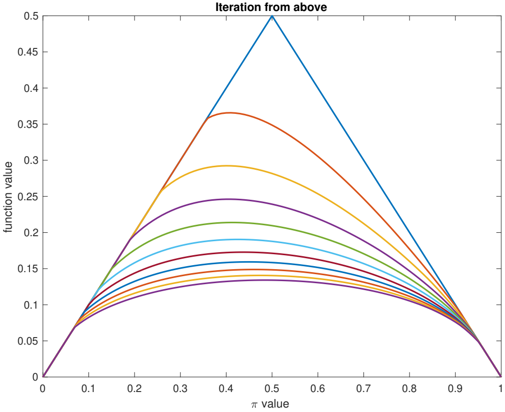
$$Tf(\pi) = \min(g(\pi), c + \mathbb{E}[f(\Pi_1)]).$$

Then we iterate:

Another important property

Claim. V is concave in π.
It suffices to show that $\mathbb{E}_\pi(f(\Pi_1))$ is concave.

Why is concavity important



Step 2: Find the optimal strategy

Define the **continuation region** and the **stopping region**:

$$C := \{\pi : V(\pi) < g(\pi)\},$$
$$D := \{\pi : V(\pi) = g(\pi)\}.$$

Then an optimal strategy is:

Some generalisations

Can we further generalise it? (1)

Continuous version (Solved in 1967)

Sequentially testing the drift of a Brownian motion:

$$X_t = \theta t + W_t.$$

Can we further generalise it? (2)

General prior distribution (Research Problem)

Instead of a Bernoulli prior, we assume a general prior μ for the unknown parameter.

Can we further generalise it? (3)

Multi-dimensional problems (Research Problem)

Nice properties depend on independence.

How much better do we achieve (Research Problem)

How much better do we get by optimising over stopping times?

$$0 \leq \bar{V} - V \leq ?$$

Thank you!