1 Experiment 1

1.1 Experimental Design

This consists of playing one round of either 3 or 15 boxes. Users win 0 the first round. Then are asked the probability of winning at least one the next round. Then they are asked if they would like to play another round, where they will keep their winnings, but they will forfeit a 1USD consolation prize they could have now.

1.2 Proof of Optimality

Let's first consider the case where one makes 3 attempts and gets a payoff of 0 each time. What can we say about the outcome of our next 3 attempts? And in the case where one makes 15 attempts and gets a payoff of 0 each time, what can we then say about the outcome of our next 15 attempts? Our estimated expected payoff is 0 in both cases, but in which case are we most confident about that expectation?

The problem is clearly a discrete 0,1 problem: either you are rejected or you are accepted. The problem of learning then becomes one of estimating the free (p) parameter of a binomial distribution, given you have sent n applications. Given your estimate, $\hat{p}$, the probability of winning at least one of the next n attempts is given by:

$$P(win) = 1 - (1 - \hat{p})^n \tag{1}$$

We will formulate the process of estimating $\hat{p}$ as one of Bayesian learning, with a Beta prior, paramaterized by α and β, on the value of p. Given n failures, the corresponding posterior Beta distribution is paramaterized by:

$$\alpha'_n = \alpha$$
$$\beta'_n = \beta + n$$

The PDF of the posterior estimate of p after n failed attempts is therefore given by:

$$P(x|n) = \frac{x^{\alpha-1}(1-x)^{\beta-1+n}\Gamma(\alpha+\beta+n)}{\Gamma(\alpha)\Gamma(\beta+n)}$$

Which can be plugged back into 1 and take expectations to give us the expected payoff:

$$\mathbb{E}[win] = \int_0^1 1 - (1-x)^n \, P(x|n)dx$$

Which is increasing in n for any valid values of α and β. Any reasonable estimate of p compatible with correct Bayesian updating will, therefore, suggest that players estimate a greater probability of winning given another n chances, after n failures, as n increases.

2 Experiment 2

2.1 Experimental Design

High boxes and low boxes, you can pick one or the other.

2.2 Proof of Optimality

The optimal strategy for the game can be formulated as follows: begin with the belief that the probability of being a high type is $\frac{1}{2}$, which you are told. Try the high box first, and afterwards update your beliefs. If you won, you know for sure you are a high type (and you also exit the game). If you lose, surely the probability of your being a high type has gone down, and as such you should update your belief downward. For now, let's assume you know the true probability of winning a high box given that you are a high type, which we will call t. Remember this is also the probability of winning a low box, regardless of type. Denote the result as k, the box type as $X_i, i \in L, H$ and the indivdual's type as $a_i, i \in L, H$:

$$P(a_H|X_H, k=0) = \frac{P(k=0|a_H, x_H)P(a_H)}{\sum_i P(k=0|x_H, a_i)}$$

Which we can rewrite in recursive form to find the posterior estimated probability of being a high type given the evidence of n failures:

$$\hat{P}_{H,n} = \frac{(1-t)}{(1-t)P(a_H) + (1-P(a_H))}\hat{P}_{H,n-1}$$

$$\hat{P}_{H,n} = \left(\frac{(1-t)}{1-P(a_H)t}\right)^n P(a_H)$$

We can then frame the game in the form of a dynamic programming problem with the following recursive value function:

$$V(x, \hat{P}_{H,n}, k_n) = \begin{cases} W(x), \text{for } k_n = 1 \\ \max\left\{\mathbb{E}[V(x_L, \hat{P}_{H,n+1}, k_{n+1})] \;,\; \mathbb{E}[V(x_H, \hat{P}_{H,n+1}, k_{n+1})]\right\}, \text{for } k_n = 0 \end{cases}$$

We can simplify the problem further if we assume, as we have, that players will start with high boxes and does not switch back after moving to low boxes. This allows us to simplify our state to n, the number of failed high-box attempts:

$$V(n, k_n) = \begin{cases} W_H, \text{for } k_n = 1 \\ \max\left\{ V_L(n) \ , \ \mathbb{E}[V(n+1, k_{n+1})] \right\}, \text{for } k_n = 0 \end{cases}$$

Where $V_L(n) = \left(1 - (1-t)^{N-n}\right) W_L$ is the expected value of the deterministic policy of playing only low boxes for the rest of the game $(\pi(n, k_n) = x_L)$. We can write out the expected value explicitly:

$$\mathbb{E}[V(n, k_n)] = t\hat{P}_{H,n}W_H + (1-t)\hat{P}_{H,n}\mathbb{E}[V(n+1, k_{n+1})] + (1-\hat{P}_{H,n})\mathbb{E}[V(n+1, k_{n+1})]$$

Intuitively, more lost attempts, given a finite number of trials, cannot be better. This means that if:

$$V_L(n) > \mathbb{E}[V(n+1, k_{n+1})]$$

this implies that:

$$V_L(n+1) > \mathbb{E}[V(n+2, k_{n+2})]$$

In other words, if we switch to low boxes, we will not want to switch back. This allows us to write the conditions for switching strategies as:

$$V_L(n) = t\hat{P}_{H,n}W_H + (1-t)\hat{P}_{H,n}V_L(n+1) + (1 - \hat{P}_{H,n})V_L(n+1)$$

Which implicitly defines n as a function of t, W_H, W_L, N, and $\hat{P}_{H,n}$, all of which are known. Plugging in $V_L(n)$ and rearranging terms, we get an interpretable version:

$$1 - (1-t)^{N-n} - (1 - t\hat{P}_{H,n})\left(1 - (1-t)^{N-n-1}\right) = t\hat{P}_{H,n}\frac{W_H}{W_L}$$

Where we can see that the expected lost value of not switching to low boxes this period must equal the expected increase in winnings from trying one more high box. Setting $t = \frac{2}{N}$ and solving for n as a function of N:

$$N = 9, n \approx 4.25$$
$$N = 45, n \approx 24.5$$

Which gives us our benchmark for optimal play. Of note, the ratio n/N is increasing in N, but only slightly.