Application of Probability: Detecting cheating in Minecraft Speedrun

Ngai Ka Shing, Sam (SID: 1155190647)
Cheung Wan Fung, Wilson (SID: 1155175377)

Abstract

In October 2020, Dream, a renowned famous Minecraft YouTubers, was accused of cheating during his numerous speedrun attempts because of "being too lucky" in two events: Piglin bartering and collecting blaze rods. Later, the Minecraft Speedrunning Team (MST) published a detailed 29-pages report concluding he cheated. This report aims to investigate the claims in the MST paper, provide evidence for such claims, and deduce what suitable modified probability Dream should use to remain unsuspicious. This report is divided into two sections: Determining the naive probability and Deducing a suitable modified probability.

1 The naive probability of getting as lucky as Dream

This section will explain why the claimed naive probability is correct in the MST paper.

1.1 Introduction

Both Piglin bartering and Blaze Rod dropping have a certain probability of obtaining desired items. Each attempt is an independent event, and we can use the binomial distribution to find out the odds of Dream.

1.2 Piglin bartering

1.2.1 Method

For each trade, there is a fixed probability of $\frac{20}{423}$ of obtaining an Ender Pearl. Considering that Dream achieved 42 Ender pearl trades out of 262 Piglin Barters, statistical modeling using Binomial(262, $\frac{20}{423}$) distribution could be carried out. By comparing Dream's results with the expected distribution, the likelihood of these results could be assessed. To evaluate this, p-value (which is the probability under null hypothesis, of obtaining a result equal to or more extreme than the observed data) could be calculated, which provides a measure to assess the likelihood of Dream's results and determine whether they are statistically significant (With p-value ≤ 0.05).

1.2.2 Code Simulation

Below are the Code simulation (10000 simulations) of Ender Pearl trade event by using Jupyter Notebook.

Submitted to 37th Conference on Neural Information Processing Systems (NeurIPS 2023). Do not distribute.

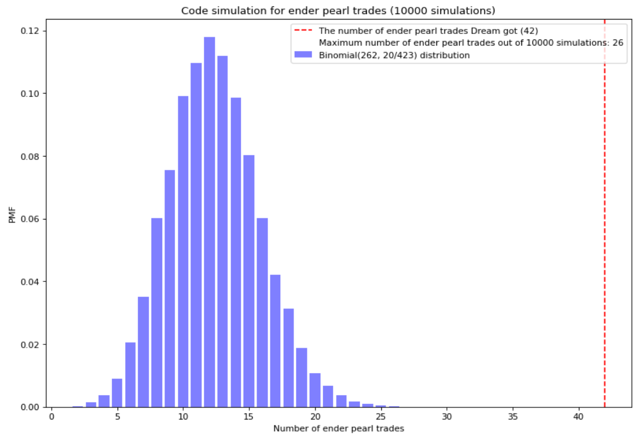

Figure 1: Binomial distribution of Ender Pearl trade event using code simulation
Link: https://github.com/sam1037/Probability-project-estr2018-/tree/main

1.2.3 Finding out the p-value

The p-value for Ender Pearl trade event is approximated as follows:

Let X be the number of Ender Pearl obtained.

$$P(X \geq 42) \approx \sum_{k=42}^{262} \binom{262}{k} \left(\frac{20}{423}\right)^k \left(1 - \frac{20}{423}\right)^{262-k} \approx 5.6 \times 10^{-12}.$$

It could be calculated that the p-value of Dream's results in the Ender Pearl trade event is $\approx 5.6 \times 10^{-12}$, which is much lower than the threshold for being classified as statistically significant.

1.3 Blaze Rods drops

1.3.1 Method

Similarly, n = 305, p = $\frac{20}{423}$ for this event. Thus the distribution would be Binomial(305, $\frac{20}{423}$). Again p-value will also be determined.

1.3.2 Code Simulation

Below are the Code simulation (10000 simulations) of Blaze Rod event by using Jupyter Notebook.

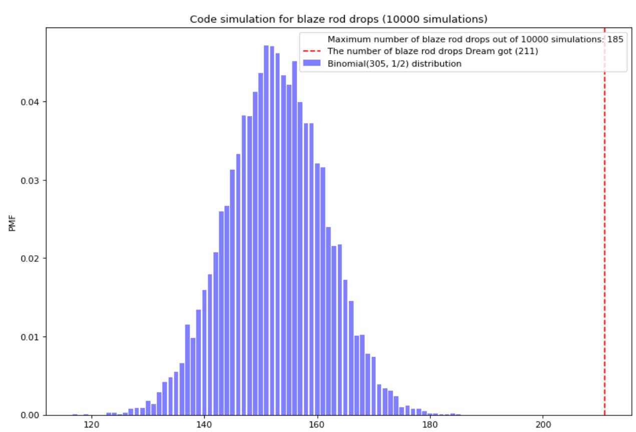

Figure 2: Binomial distribution of Blaze Rod event using code simulation
Link: https://github.com/sam1037/Probability-project-estr2018-/tree/main

1.3.3 Finding out the p-value

The p-value for Blaze Rod event is approximated as follows:

Let X be the number of Blaze Rod obtained.

$$P(X \geq 211) \approx \sum_{k=211}^{305} \binom{305}{k} \left(\frac{1}{2}\right)^k \left(1 - \frac{1}{2}\right)^{305-k} \approx 8.8 \times 10^{-12}.$$

The p-value is approximately equal to 8.8×10^{-12}, which is also much lower than the threshold for being classified as statistically significant.

1.4 Combined probability

In Dream's case, where both of the two independent events occur simultaneously, the combined probability would be equal to:

$$\begin{aligned}
&P(\text{Getting 211 Blaze Rods out of 305 trials}) \times P(\text{Getting 42 ender pearl trades out of 262 Piglin Barters}) \\
&= (8.8 \times 10^{-12}) \times (5.6 \times 10^{-12}) \\
&\approx 5.0 \times 10^{-23},
\end{aligned}$$

which is almost equivalent to being struck by lightning for 3.56×10^{16} consecutive days, indicating that it is reasonable to conclude that it is impossible.

2 Deduce a suitable modified probability

This section aims to deduce a suitable modified probability that Dream should use to remain unsuspicious most of the time. We would apply the Central Limit Theorem to approximate both binomial distributions as normal distribution.

Method: First, we establish a threshold and assume values less than or equal to that threshold are considered unsuspicious. In this case we establish a lenient threshold of the **mean plus 3 standard deviation**, that is, roughly $\Phi(3) = 99.87\%$ of the unmodified distribution are being considered unsuspicious. Then we find a suitable modified probability such that at least **95%** of the modified distribution are unsuspicious.

2.1 Blaze rods

We will calculate for blaze rod first as the numbers are nicer. Recall that the original probability of a Blaze dropping any Blaze Rods is 0.5, and Dream killed 305 Blazes. So p is 0.5 and n is 305. We let **X** be a Binomial(305, 0.5) random variable, which represents the unmodified distribution of the number of blazes dropping blaze rod(s). We can find the threshold t as follows:

$$\begin{aligned}
\mu_x &= E[X] = 305 \times 0.5 = 152.5 \\
\sigma_x^2 &= Var[X] = 305 \times 0.5 \times (1 - 0.5) = 76.25 \\
\sigma_x &= \sqrt{Var[X]} = \sqrt{76.25} \approx 8.732124598 \\
t &= \mu_x + 3\sigma_x \approx 178.6963738
\end{aligned} \tag{1}$$

Now we let **Y** be a Binomial(305, 0.5m) random variable representing modified distribution, where m denotes the modifying constant that increases Dream's luck, which is greater than or equal to 1, and the 0.5 comes from the unmodified probability. Note that mean $\mu_y = 152.5m$, and variance $\sigma_y^2 = 152.5m(1 - 0.5m)$. We want at least 95% of the modified distribution to remain unsuspicious, so we want to solve m for the following:

$$P(Y \leq t) \geq 0.95 \tag{2}$$

The L.H.S. of 2 is:

$$\begin{aligned}
&P(Y \leq t) \\
=&P(\frac{Y-\mu_y}{\sigma_y} \leq \frac{t-\mu_y}{\sigma_y}) \\
\approx&\Phi(\frac{t-\mu_y}{\sigma_y})
\end{aligned} \tag{3}$$

where 3 is by the Central Limit Theorem

Since Φ is a increasing function, and $0.95 \approx \Phi(1.644853627)$, by 2 and 3 we have:

$$\begin{aligned}
\frac{t-\mu_y}{\sigma_y} &\geq 1.644853627 \\
t^2 - 2t\mu_y + \mu_y^2 &\geq 1.644853627^2 \cdot \sigma_y^2
\end{aligned} \tag{4}$$

Solving the inequality in 4, we have:

$$m \leq 1.077881528 \text{ or } m \geq 1.262656682 \text{ (rej. since } t - \mu_y \geq 0)$$

Dream would like to have a greatest possible m, so the modify constant $\boldsymbol{m \approx 1.077881528}$. Therefore, **Y** is a Binomial(305, 1.077881528*0.5) random variable, with mean $\mu_y \approx 164.3769331$ and variance $\sigma_y^2 \approx 75.78750315$.

Thus, we can see that if Dream were to use a conservative modifying constant, he would on average only get $m - 1 \approx 7.79\%$ more blazes to drop blazes rod(s), which would not give him a substantial advantage. **Therefore, with the knowledge of probability, one could conclude that Dream could not get a substantial advantage while being unsuspicious in this event.**

A simulation was ran for the modified distribution.

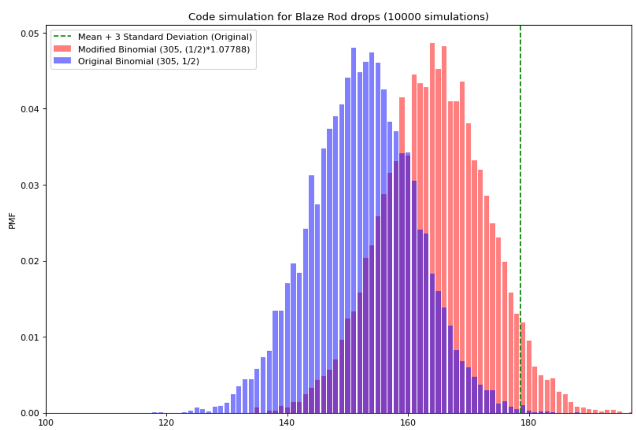

Figure 3: Blaze Rod event with modified probability
Link: https://github.com/sam1037/Probability-project-estr2018-/tree/main

The p-value for the modified distribution is approximated as follows:

Let X be the number of Blaze Rod drops.

$P(X \geq 178.6963738) \approx \sum_{k=179}^{305} \binom{305}{k} \left(\frac{1}{2} * 1.077881528\right)^k \left(1 - \frac{1}{2} * 1.077881528\right)^{305-k} \approx 0.049737257 \leq 0.05$,

And it supports our claims.

2.2 Ender Pearl trade

Similarly, in the Ender Pearl trade event, p = 20/423 and n = 262. Take X = Binomial(262, 20/423). We can find the threshold t as follows:

$$\begin{aligned}
\mu_x &= E[X] = 262 \times \frac{20}{423} = 12.38770686 \\
\sigma_x^2 &= Var[X] = 262 \times \frac{20}{423} \times (1 - \frac{20}{423}) = 11.80199968 \\
\sigma_x &= \sqrt{Var[X]} = \sqrt{11.802} \approx 3.435403859 \\
t &= \mu_x + 3\sigma_x \approx 22.69391844
\end{aligned} \tag{5}$$

Again, let Y = Binomial(262, $\frac{20}{423}$m), which represents modified distribution.Then mean $\mu_y = 12.38770686m$, and variance $\sigma_y^2 = 12.38770686m(1 - \frac{20}{423}m)$. We would like to solve the following:

$$P(Y \leq t) \geq 0.95 \tag{6}$$

Using the same technique in section 2.1, we have:

$$m \leq 1.313312171 \text{ or } m \geq 2.529316993 \text{ (rej. since } t - \mu_y \geq 0)$$

Take m = 1.313312171. Therefore, **Y** is a Binomial(262, 0.062095138) random variable, with mean $\mu_y \approx 16.26892619$ and variance $\sigma_y^2 \approx 15.25870497$.

Thus, if Dream were to use a conservative modify constant, he would on average only get $m - 1 \approx 31.3\%$ more blazes to drop blazes rod(s), which would give him a substantial advantage. **Therefore, with the knowledge of probability, one could conclude that Dream could get a substantial advantage while being unsuspicious in this event.**

A simulation was ran for the modified distribution.

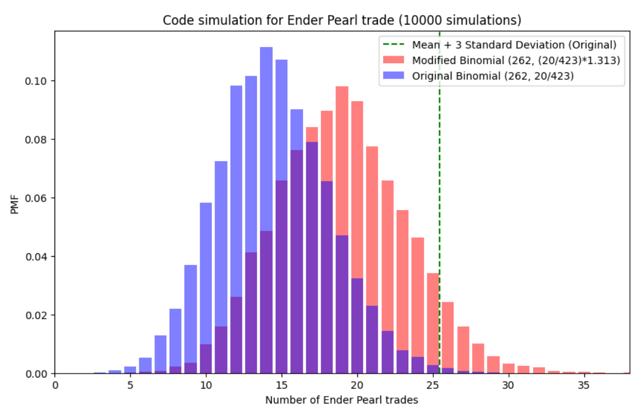

Figure 4: Ender Pearl trade event with modified probability
Link: https://github.com/sam1037/Probability-project-estr2018-/tree/main

The p-value for the modified distribution is approximated as follows:

Let X be the number of Ender Pearl obtained.

$P(X \geq 22.69391844) \approx \sum_{k=23}^{262} \binom{262}{k} \left(\frac{20}{423} * 1.313312171\right)^k \left(1 - \frac{20}{423} * 1.313312171\right)^{262-k} \approx 0.0498663781 \leq 0.05$,

And it supports our claims.

3 References

Minecraft Speedrunning Team. "Dream Investigation Results Official Report." Minecraft Speedrunning Team, 11 Dec. 2020. Updated 15 Dec. 2020.