1. (the Exchange Paradox)
   1. Given we know that the amount of money put in envelope 1 is m, then the probability of you finding m or 2m dollars in your envelope is X, so given we know m we can write the probability of X as:

Now if we try to determine the probability of M’s value based on X we can describe that as:

We can use this too write the probability of getting x vs when trading, given you know X = x, as the odds between the two outcomes:

Given all this we will let Y stand for the amount of money in our opponent’s envelope, where . So we need an equation for the expected value of Y given you know X:

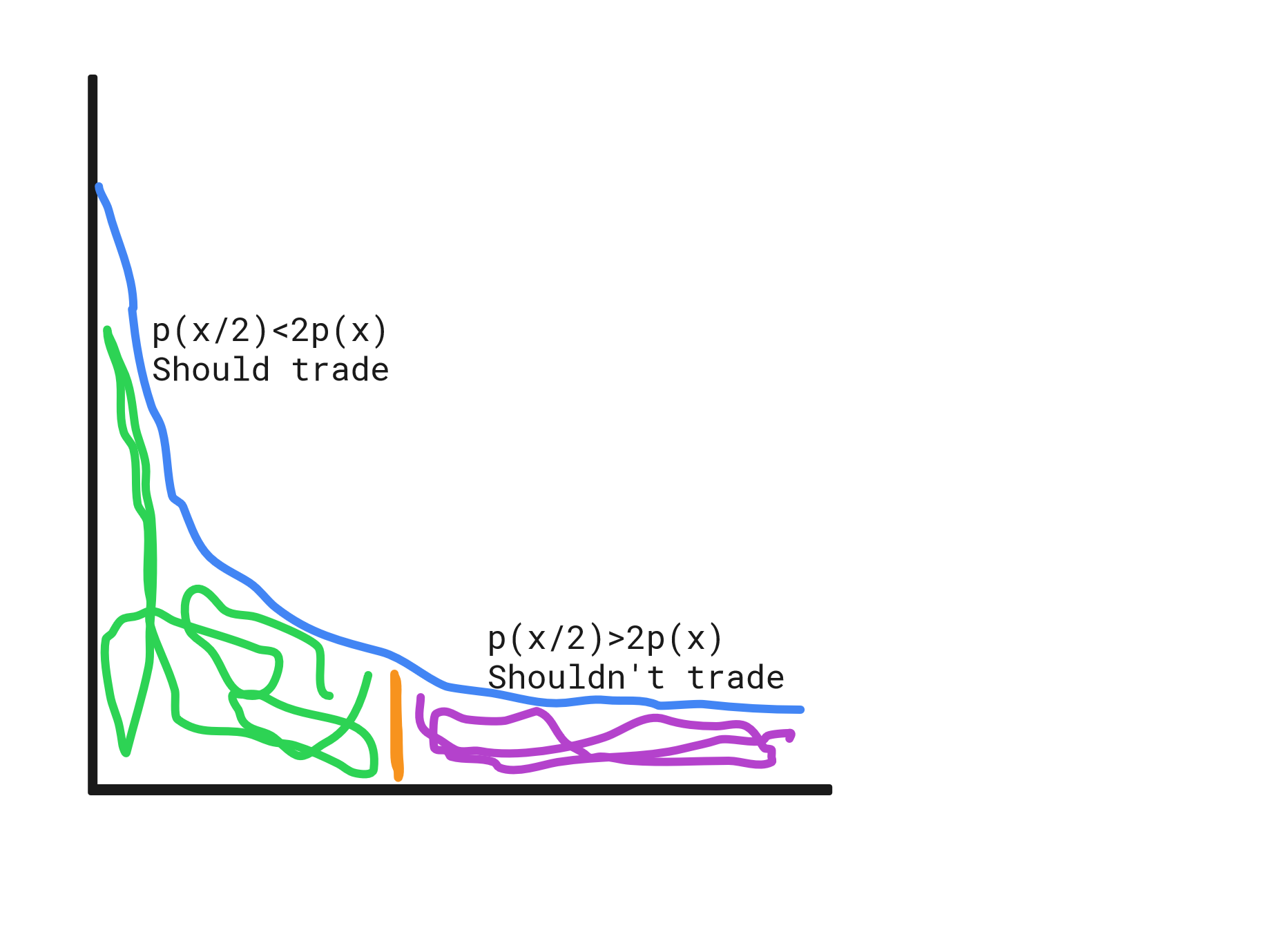
* 1. In this scenario the action space has two possible actions :

We will define the utility function as:

We will now define the expected value function for each possible action:

If then you should offer to trade, however if this isn’t true then you shouldn’t offer to trade, we will now solve the inequality:

given this you should only trade if and only if: .

Drawing of when you should and shouldn’t trade using an exponential model:

If your model for m is mExponential () and you see x dollars in your envelope you should offer to trade if and only if:

. For this prior, the optimal action is to offer to trade if and only if the amount of money in your envelope is x and the amount of money in envelope 1 is E(M) and the following holds given this: .

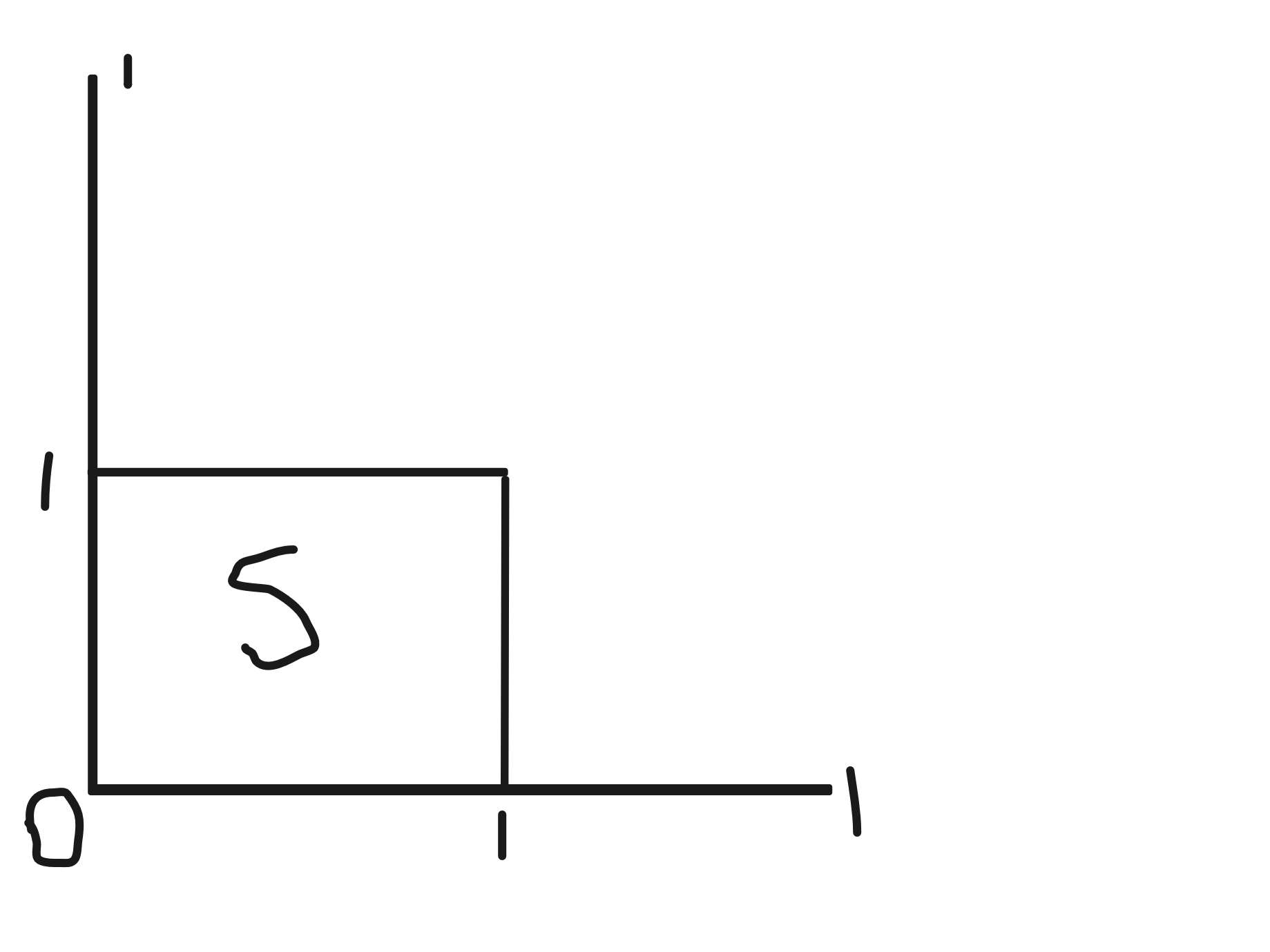
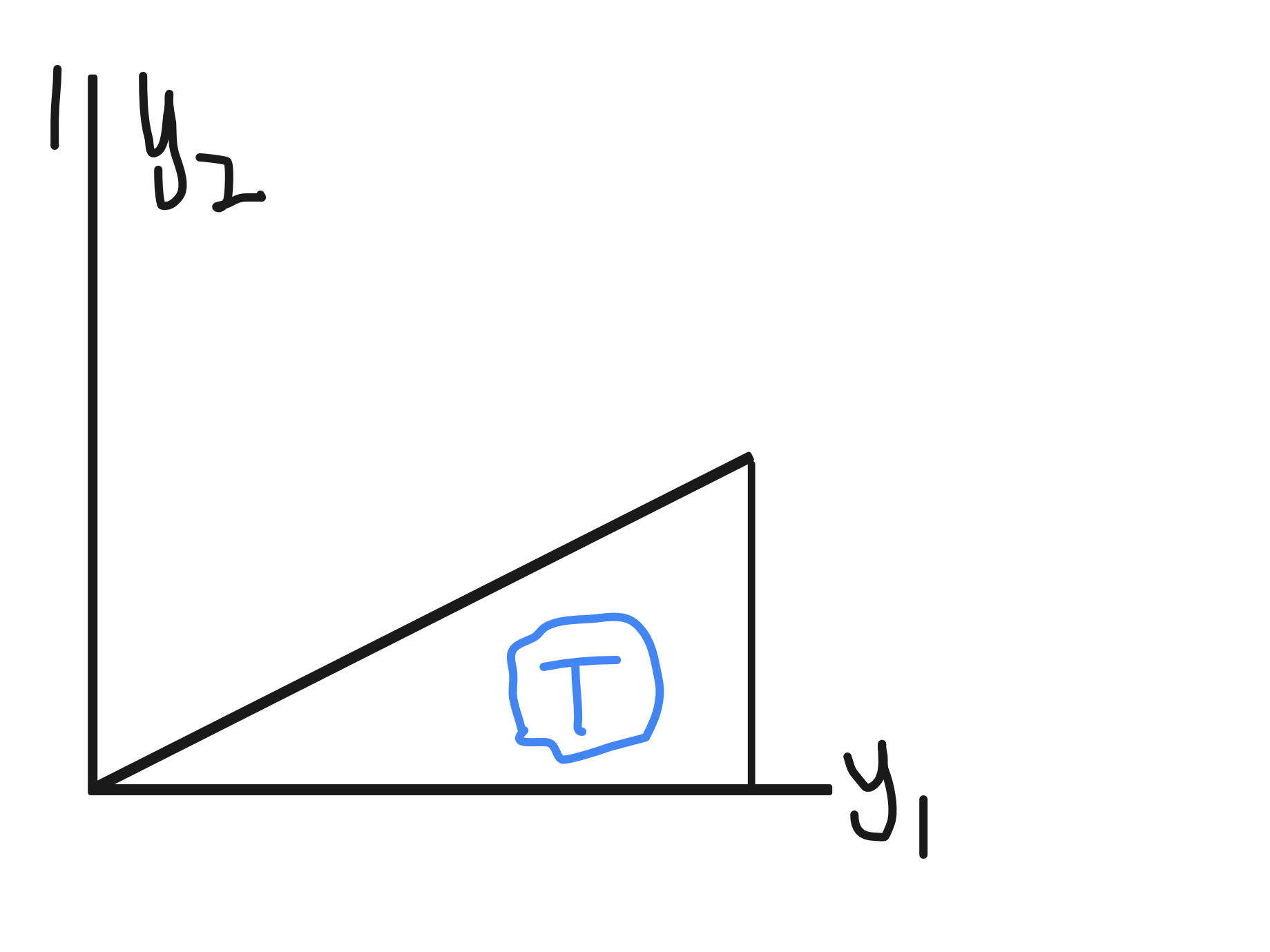
* 1. The incorrect belief in paragraph 2 is that rather than calculating the proper conditional probabilities for the expected value equation, they just used for both monetary outcomes giving them an incorrect result of . However this means they did not account for the new information gained when you see how much money is in your envelope, which plays a very big role in determining the actual expected value.

1. (practice with joint, marginal and conditional densities)
   1. The marginal for is:

The marginal for is:

The product of these two marginals is:

Since the product equals the joint PDF must be independent.

* 1. The correlation must be zero as the two variables are independent.
  2. 

We will now verify that the joint density function integrates to 1: .

* + 1. The marginal distribution for is:

The marginal distribution for is:

Choose a fixed y2 between 0 and y1 and now integrate over y1:

Choose a fixed y1 between 0 and 1 and now integrate over y2:

* + 1. Fix such that ,

Fix such that ,

* + 1. Given we already know the conditional expectations and how to calculate the variance of a random variable we just need to compute these values:
  1. They are dependent for two reasons, first because the support requires y2 is less than y1 and second if you multiply the two marginals distributions you don’t get the joint distribution:
  2. In order to calculate the correlation of we first need to find the standard deviation and expected value of each:

1. (moment-generating functions)
   1. We know that the function for the skewness of x is: , we also know that: using all this we can now better define the skewness of x as: and this will eventually equal: . The skewness will vanish when as in that case the numerator of the skew will be 0, thus making it 0. As n goes toward infinity the skewness tends toward 0, as the denominator will eventually dwarf the numerator.
      1. We know that and the MGF of a discrete random variable is: . So given all this the MGF for the discrete random variable Y is: now we have to use the knowledge that and say that giving us the MGF:
      2. The first moment is:

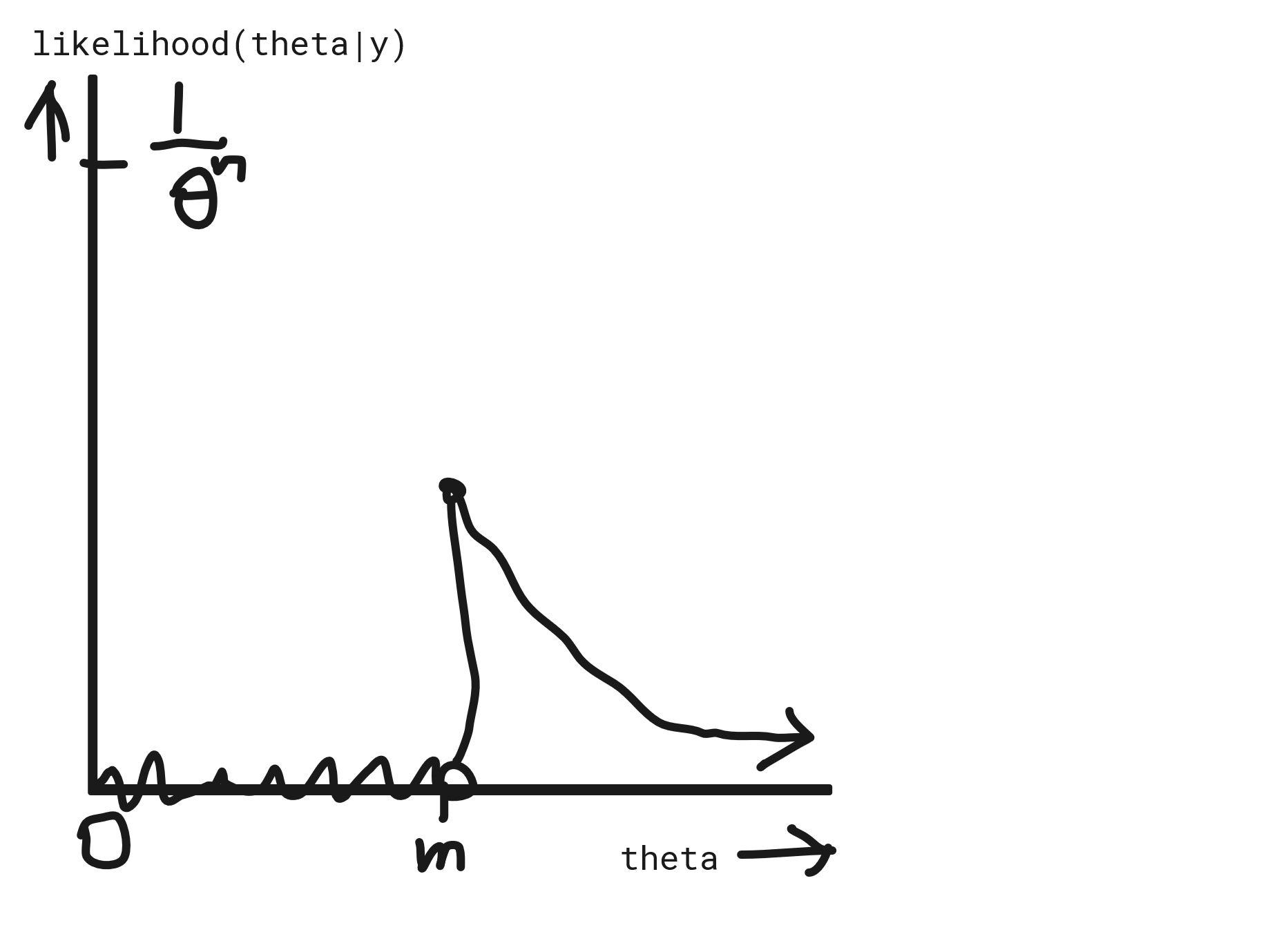
The second moment is:

The third moment is:

The variance of Y is:

Since Y is a Poisson distribution its skewness is:

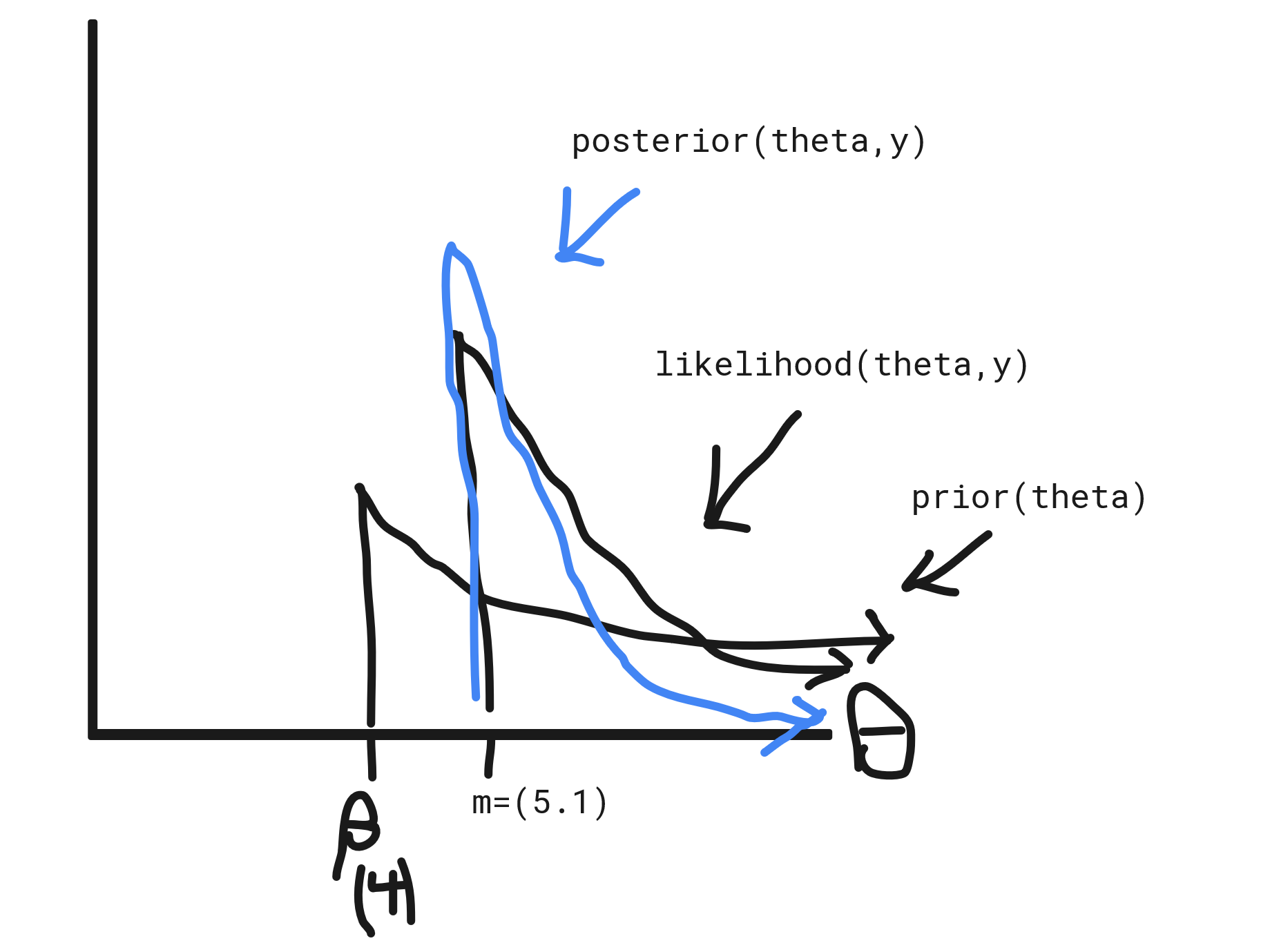
The skewness tends to zero as approaches .

1. (archaeology)
   1. Since the model is uniform and IID, the marginal distribution for each , because of this as the multiplying of all the indicators will only be 1 if the max y value is less than , so in essence this is the only thing that matters for evaluating if the indicator will return true when determining the joint distribution. Since the model is uniform and IID its joint distribution is just the product of all marginals:
      1. 

Based on this graph the MLE for this problem must be: as the likelihood function has its greatest value at m.

* + 1. The function is discontinuous at its maximum so it can’t be differentiated to find the maximum value.
    2. It is normalized if you use the c in the likelihood function to ensure it integrates to 1, doing this allows you to model this function using a more standard PDF. Given this we can model the likelihood as the pareto distribution with the parameters , as when the indicator function is the same for both when you treatas w, additionally we can replace with c as this is just for normalization, and finally we can make when .
    3. If we take both the prior distribution (parameters: ) and the likelihood function (parameters: ) as pareto distributions, then the resulting posterior distribution is just another pareto distribution as any two pareto distributions multiplied together is another pareto distribution. The math showing this can be found below: .

Therefore our resulting posterior distribution is a pareto distribution with parameters: .

* + 1. In the provided data set m = 5.1 and n = 11. We are also given that using this we get the below graph:
    2. Given our values provided in (i) we can now calculate the posterior mean and SD below:

Posterior variance: given this the posterior SD is: and the mean is: . We can now rewrite the sentence from the problem: *On the basis of this prior and data information, the value for this species of fossil ammonite is about , give or take about* .