

Exercise (Independence of biased coin throws). Consider an experiment where we first choose a biased coin from the box and then throw the selected coin four times. The box contains two types of coins: tail light and tail heavy coins. The probability of heads is 90% for a tail light coins and 10% for tail heavy coins. The probability of choosing a tail light coin is unknown. Compute the conditional probability $\Pr[\text{head}_2|\text{head}_1]$ and explain whether two throws of the coin are independent or not. Are two throws of the coin independent if we know whether the coin is tail light or tail heavy? Compute the probability that all four coins are heads. Estimate the probability of the same event from below if we do not know anything about the coins. Instead, we know that if we choose a coin and throw it then the probability of obtaining head is 50%.

Solution. Let p denote the probability that we get a tail-light coin out of the box. Then the probability that the first coin lands on heads is

$$\Pr[\text{head}_1] = p \cdot 0.9 + (1 - p) \cdot 0.1 = 0.8p + 0.1 \ .$$

At the same time, the probability of obtaining two heads in the row is

$$\Pr[\text{head}_1 \wedge \text{head}_2] = p \cdot 0.9 \cdot 0.9 + (1 - p) \cdot 0.1 \cdot 0.1 = 0.8p + 0.01 \ .$$

Therefore, we can express the conditional probability as follows:

$$\Pr[\text{head}_2|\text{head}_1] = \frac{\Pr[\text{head}_1 \wedge \text{head}_2]}{\Pr[\text{head}_1]} = \frac{0.8p + 0.01}{0.8p + 0.1} = 1 - \frac{0.09}{0.8p + 0.1} \ .$$

By using the total probability formula and event decomposition, we can estimate

$$\Pr[\text{head}_2] = p \cdot 0.9 \cdot 0.9 + p \cdot 0.1 \cdot 0.9 + (1 - p) \cdot 0.1 \cdot 0.1 + (1 - p) \cdot 0.9 \cdot 0.1 = 0.8p + 0.1 \ .$$

It is easy to see that the first and the second coin toss are not independent, as the equation $\Pr[\text{head}_2] = \Pr[\text{head}_2|\text{head}_1]$ does not hold in general. To study this phenomenon more deeply, let us study the ratio

$$\gamma = \frac{\Pr[\text{head}_2|\text{head}_1]}{\Pr[\text{head}_2]} = \frac{0.8p + 0.01}{(0.8p + 0.1)^2} = 1 + \frac{0.64p(1 - p)}{(0.8p + 0.1)^2} \ .$$

From the form it is clear that $\gamma > 1$, which means that probability of obtaining two heads is always larger than the probability of two heads if the throws would be independent, see Figure 1. Since initial conditions are symmetrical for heads and tails, the same reasoning shows that the probability of obtaining two tails is also overrepresented compared to the independent model. As a result, probabilities of obtaining different outcomes must be smaller than under independence assumption. The latter is an expected result, as the selected coin is strongly biased we should get the preferred value with higher probability.

Now let us assume that we know which coin was chosen. Without loss of generality we can assume that we chose a tail-light coin. Then by definition

$$\Pr[\text{head}_1|\text{tail-light}] = \Pr[\text{head}_2|\text{tail-light}] = 0.9$$

and by description of the experiment

$$\Pr[\text{head}_1 \wedge \text{head}_2|\text{tail-light}] = \Pr[\text{head}_1|\text{tail-light}] \Pr[\text{head}_2|\text{tail-light}] = 0.81 \ .$$

Hence, the coin-tosses are indeed independent if we know the type of the coin. Note that for $p = 1$ we know that the tail-light coin is always chosen and thus the over-representation coefficient $\gamma(1) = 1$. The same reasoning also gives that $\gamma(0) = 1$. These conditions are indeed met, see Figure 1.

The probability that in four throws all coins land on the head is

$$\Pr[\text{head}_1 \wedge \dots \wedge \text{head}_4] = p \cdot 0.9^4 + (1 - p) \cdot 0.1^4 \ .$$

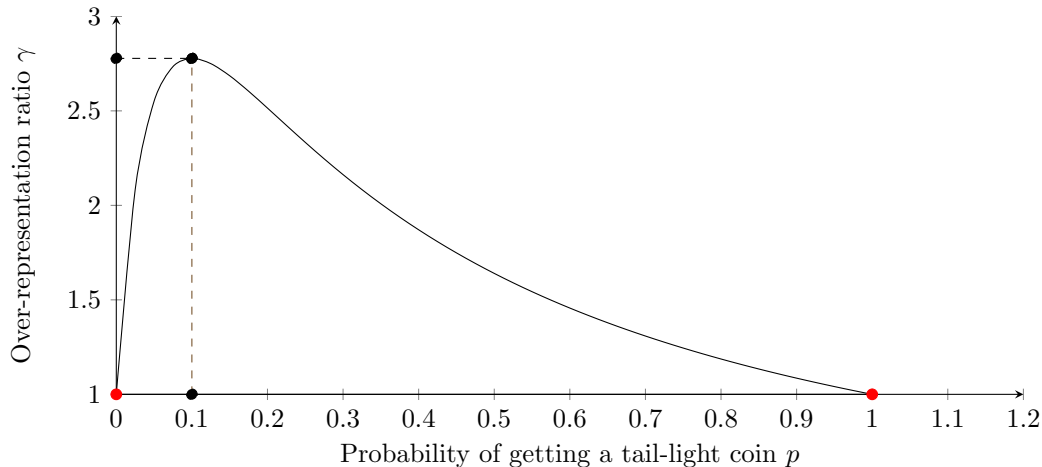


Figure 1: Over-representation ratio γ as a function of selection probability p . The ratio decreases when the choice of tail light coin becomes more certain and becomes one if the choice is certain.

For example, $\Pr[\text{head}_1 \wedge \dots \wedge \text{head}_4] = 0.3281$ if both coin types are equiprobable: $p = 0.5$.

BOX WITH UNKNOWN COINS. In this setting we do not know what kind of coins are in the box. We just know that if we take a random coin then the probability of getting a head as an outcome is 50%. Let us assume that there are n different types of coins x_1, \dots, x_n that occur with probabilities p_1, \dots, p_n . Also assume that the coin x_i lands on head with probability λ_i . Then

$$\Pr[\text{head}_1] = \sum_{i=1}^n p_i \lambda_i = 0.5$$

by assumption. Now we can express the probability of for heads as follows:

$$\Pr[\text{head}_1 \wedge \dots \wedge \text{head}_4] = \sum_{i=1}^n \Pr[x_i \wedge \text{head}_1 \wedge \dots \wedge \text{head}_4] = \sum_{i=1}^n p_i \lambda_i^4 .$$

Note that this holds, as the trows become independent if we know the type of coin. As x^4 is a convex-cup function, Jensen's inequality assures that

$$\Pr[\text{head}_1 \wedge \dots \wedge \text{head}_4] = \sum_{i=1}^n p_i \lambda_i^4 \geq \left(\sum_{i=1}^n p_i \cdot \lambda_i \right)^4 = 0.5^4 = 0.0625 .$$

This means that the probability throwing four heads in a row is at least 0.0625 given that the probability of throwing one head is 0.5. The bound is quite loose in general. Above, we established that if tail-heavy and tail-light coins are equiprobable then $\Pr[\text{head}_1] = 0.5$ and $\Pr[\text{head}_1 \wedge \dots \wedge \text{head}_4] = 0.3281$, which is much larger than our lower bound 0.0625.