Days ago I was surfing online and found a simple question like this: if you toss a coin 20 times and had 13 head and 7 tails, can you make any judgements about whether the coin is a fair coin? At first I didn't have much idea of how to solve this problem. After searching and reading some articles I knew the answer.

There were at least 3 methods this question could be solved: (1) using normal distribution Z score. (2) Using probability to find the p value/type 1 error. (3) Using chi-square goodness of fit test. All methods yield same answer.

Using normal distribution.

As we know, if the coin is fair, toss it for 20 times, on average, you will get 10 heads and 10 tails. Let x=1 when head appears and let x=0 when tail appears. Then x is a 0-1 distribution in any one toss. If we toss n times, the number of head or tail becomes a binomial distribution. According to binomial distribution, if x~B(n,p), p is the probability of event A to happen, then the expectation of number of event A to happen is n*p. In our case, let A = head appears, p=0.5, therefore on average event A (head appears in one toss) happens 20*0.5=10 times in 20 tosses. We can also find variance of number of event A is np*(1-p) =20*0.5*0.5=5. Std dev=sqrt(5)

In our case, since number of head is 13 (x=13), we can find how much deviation our example is from the mean: 13-10=3. We can find the Z score is: (13-10)/sqrt(5)=1.342. That means in our case, the actual result is 1.342 times of the standard deviation from the mean. Then we check the standard normal distribution table to see if this situation is rare enough that we can say the coin must be not fair to get such an unusual result. The Z score for 0.05 significant level for two tail is 1.96. Therefore, we can say that our case is not "unusual enough" to conclude that the coin is unfair.

If we have 14 head, then Z score will be (14-10)/sqrt(5)=1.789, still less than 1.96. If we have 15 head, then Z score will be (15-10)/sqrt(5)=2.236, bigger than 1.96! Then we can say that the coin is very unusual to be a fair coin because it give us such an extreme Z score, therefore, we will conclude that the coin is not fair! Please note, there is still a chance that you make the wrong conclusion. But out of 100 times, you make 95 correct conclusions. (Alpha=0.05).

Using Probability.

Another method you can use is to calculate the probability directly to see if the sum of probabilities of events as least as unusual as you currently get is smaller than level of significance (0.05 in this case). In this case, we saw 13 head, and we need to sum the probability of seeing 13,14,15,…,20 head to compare with 0.05!

Probability (13 head) = 20 nCr 13 (p) ^13 *(1-p) ^7

Probability (14 head) = 20 nCr 14 (p) ^14 *(1-p) ^6

……

Probability (20 head) = 20 nCr 0 (p) ^0 * (1-p) ^20

Adding the left items=adding the right items= (0.5^20)*(20 nCr13+ 20 nCr14+ 20 nCr15+…+20 nCr20) =0.132. As 0.132>0.05 we can not reject null hypothesis that the sample is not from the population.

Assume we have 14 head, then the sum of probability is:

(0.5^20)*(20 nCr14+ 20 nCr15+…+20 nCr20) =0.058>0.05, so we still can not reject.

Assume we have 15 head, then the sum of probability is:

(0.5^20)*(20 nCr15+…+20 nCr20) =0.021<0.05, so we can reject the null hypothesis and say that the coin is not fair!

Using Chi-square goodness of fit test.

According to chi-square goodness of fit theorem, we can use the following formula to test if the sample data set is from the population. The formula is:

$$\chi^2 = \sum_{i=1}^{r} \frac{(m_i - n * p_i)^2}{n * p_i}$$ with degree of freedom of (r-1).

m_i is the number of times event A_i happened in your sample. r is the number of distinct result (in toss coin example, r is 2, head or tail). p_i is the theoretical probability of event A_i to happen. $\sum_{i=1}^{r} p_i$ =1.

In our case, r=2, m1=13, m2=7, $n * p_i$ =20*0.5=10.

Therefore, $\chi^2 = \sum_{i=1}^{2} \frac{(m_i - 10)^2}{10}$ =1.8, degree of freedom is 2-1=1.

Use χ^2 table we find χ^2 value with 1 degree of freedom and alpha=0.05 to be 3.841, which is bigger than 1.8. Therefore, our conclusion is that the sample is not unusual enough to make the coin an unfair one.

Assume we saw 15 head and 5 tail, then we recalculate χ^2, which equals 5, bigger than 3.841. Therefore, we conclude that the coin is unfair. The conclusion is the same as before.