MthStat 465, Spring 2005, Lecture Number 18
A simple example of testing hypotheses

Suppose I have two coins. Coin 1 comes up heads with probabiity 1/2 and coin 2 comes up heads with probability 1/3. I choose one coin at random, and I ask you to determine which coin it is by tossing the coin 5 times, and using only the number of heads obtained. You need to come up with a rule that makes a decision. Once you have the rule, we are interested in two quantities:

- What is the probability that your rule decides you have coin 1 when you have coin 1;
- What is the probablity that your rule decides you have coin 2 when you have coin 2.

Your rule can be thought of as a random variable taking value 1 if you choose coin 1 and 2 if you choose coin 2. You have a sample space $\{0, 1, 2, 3, 4, 5\}$ with each element the number of heads, you have all subsets as the set of events, but you have two possible assigments of probabilities.

For example, suppose your rule was $R(s) = 1$ if $s \in \{2, 3\}$ and $R(s) = 2$ if $s \in \{0, 1, 4, 5\}$. Then if you indeed have coin 1 the probability you make the right decision is

$$\binom{5}{2}\left(\frac{1}{2}\right)^5 + \binom{5}{3}\left(\frac{1}{2}\right)^5 = \frac{20}{32} = 0.625,$$

while if you have coin 2 the probability of a correct decision is

$$\begin{aligned} & \binom{5}{0}\left(\frac{1}{3}\right)^0\left(\frac{2}{3}\right)^5 + \binom{5}{1}\left(\frac{1}{3}\right)^1\left(\frac{2}{3}\right)^4 \\ & + \binom{5}{4}\left(\frac{1}{3}\right)^4\left(\frac{2}{3}\right)^1 + \binom{5}{5}\left(\frac{1}{3}\right)^5\left(\frac{2}{3}\right)^0 \\ = & \frac{123}{243} \approx 0.506. \end{aligned}$$

Neither result is very encouraging, as you are barely right half the time.

We could try a different rule: $R(s) = 1$ if $s \in \{3, 4, 5\}$ and $R(s) = 2$ if $s \in \{0, 1, 2\}$. Then if you indeed have coin 1 the probability you make the right decision is

$$\binom{5}{3}\left(\frac{1}{2}\right)^5 + \binom{5}{4}\left(\frac{1}{2}\right)^5 + \binom{5}{5}\left(\frac{1}{2}\right)^5 = \frac{16}{32} = 0.500,$$

while if you have coin 2 the probability of a correct decision is

$$\binom{5}{0}\left(\frac{1}{3}\right)^0\left(\frac{2}{3}\right)^5 + \binom{5}{1}\left(\frac{1}{3}\right)^1\left(\frac{2}{3}\right)^4 + \binom{5}{2}\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^3 = \frac{192}{243} \approx 0.790.$$

This is pretty encouraging if you hold coin 2, but not if you hold coin 1.

The solution lies in tossing the coin more. The underlying idea is that if H is the number of heads in N tosses, if H/N is near 1/3 we think we have coin 2 while if H/N is near 1/2 we think we have coin 1. The variance of H/N is $p(1-p)/N$, which shrinks as N increases, so more tosses ought to give us better results. The Central Limit Theorem can help us make this precise. It seems reasonable to have our decision rule of the following form. Pick a number C and in $H < C$ then we have coin 2 while if $H > C$ we have coin 1. As long as C is not an integer, the event $H = C$ is the empty set. C should be some number between $N/3$ and $N/2$.

Here is how we can be sure of being right with probability 0.95 no matter which coin we have.

Case 1: Suppose we have coin 1, and we toss N times. We want

$$\begin{aligned} 0.95 &= \Pr(H > C) \\ &= \Pr\left(\frac{H-(N/2)}{\sqrt{N/4}} > \frac{C-(N/2)}{\sqrt{N/4}}\right) \\ &\approx \frac{1}{\sqrt{2\pi}}\int_{(C-(N/2))/(\sqrt{N/4})}^{\infty} \exp(-u^2/2)\,du \end{aligned}$$

where the approximation comes from the Central Limit Theorem. We know (for example, from a table) that

$$\frac{1}{\sqrt{2\pi}}\int_{-1.645}^{\infty} \exp(-u^2/2)\,du \approx 0.95002$$

so we should have

$$\frac{C-(N/2)}{\sqrt{N/4}} = -1.645. \tag{1}$$

Case 2: Suppose we have coin 2, and we toss N times. We want

$$\begin{aligned} 0.95 &= \Pr(H < C) \\ &= \Pr\left(\frac{H-(N/3)}{\sqrt{2N/9}} < \frac{C-(N/3)}{\sqrt{2N/9}}\right) \\ &\approx \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{(C-(N/3))/(\sqrt{2N/9})} \exp(-u^2/2)\,du \end{aligned}$$

where the approximation comes from the Central Limit Theorem. We know (for example, from a table) that

$$\frac{1}{\sqrt{2\pi}}\int_{1.645}^{-\infty} \exp(-u^2/2)\,du \approx 0.95002$$

so we should have

$$\frac{C-(N/3)}{\sqrt{2N/9}} = 1.645. \tag{2}$$

We don't know N or C, but we have (1) and (2) that we can solve together to find them. Ignoring (temporarily) that N must be an integer we get $N \approx 91.92519206$ and $C = 38.07666128$. We then round N up to 92, and use (1) and (2) to get two values of C: $C = 38.11085714$ and $C = 38.10462188$. Since they are both bigger than 38 and less than 39 our rule is to toss the coin 92 times, and decide coin 2 if there are 38 or fewer heads, and coin 1 if there are 39 or more heads.

As a final check (because of all the approximations)

$$\begin{aligned} \sum_{k=0}^{38} \binom{92}{k}\left(\frac{1}{3}\right)^k\left(\frac{2}{3}\right)^{92-k} &\approx 0.9565692704 \\ \sum_{k=39}^{92} \binom{92}{k}\left(\frac{1}{2}\right)^k\left(\frac{1}{2}\right)^{92-k} &\approx 0.9413149206 \end{aligned}$$

so we can see that we are pretty close to what we wanted. To get exactly what we want we have to increase N and change C. For example, $N = 100$, and choose coin 2 if the number of heads is less than or equal to 42, and choose coin 1 if the number of heads is 43 or more:

$$\begin{aligned}
\sum_{k=0}^{41} \binom{100}{k} \left(\frac{1}{3}\right)^k \left(\frac{2}{3}\right)^{100-k} &\approx 0.9566285096 \\
\sum_{k=42}^{100} \binom{100}{k} \left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^{100-k} &\approx 0.9556869599.
\end{aligned}$$