BCMS talk	The Dowry Problem	Oct 9, 2013

A king must select a groom for his daughter, the princess. Outside the castle 100 suitors line up, each one prepared to present his dowry in exchange for the princess's hand in marriage. The king sees the suitors one by one. Each announces his dowry, and the king must decide on the spot whether to select that suitor for his daughter or to lop off the man's head (he is a vicious and irrational ruler). Once the king's selection is made, the other suitors are sent off to the guillotine; thus, the king's selection is final. The king hopes to select the suitor with the largest dowry, but he has no information about the distribution of the dowries, nor, crucially, where in the sequence the suitor with the largest dowry appears. You are the king's wise wo/man; how do you advise the king to proceed so as to optimize the likelihood that the suitor with the highest dowry gets selected?

A naïve strategy would be to select the first suitor. The likelihood of success in this case is 1/100. A variation on this would be to select the last suitor. That way the king at least sees who had the largest dowry, but again, the likelihood of success is the same: 1/100. Indeed, any strategy which says to select the k^{th} suitor, regardless of the proposals, has likelihood of success equal to 1/100. Is there any strategy that improves upon a probability of 1%?

Indeed, there is. Consider the following strategy. The king lets the first fifty suitors pass in sequence, hearing the dowry of each. He then selects the next suitor whose dowry exceeds that of all the first fifty suitors. Of course, this strategy will fail if the suitor with the largest dowry occurs amongst the first fifty. But notice that it will succeed if the suitor with the largest dowry occurs amongst the second batch of fifty and the suitor with the second-largest dowry occurs amongst the first fifty. The likelihood of this scenario is roughly 1/4: actually, it is $(50/100) \times (50/99) > 1/4$, or more than 25%: a great improvement over 1%! Furthermore, there are other ways that the strategy could succeed or fail: for instance, if the suitor with the third-largest dowry is amongst the first 50, and the suitors with the largest and second-largest dowries are amongst the second 50, then it will succeed or fail depending on whether the suitor with the largest dowry precedes or follows the one with the second-largest.

In fact, we can determine the probability of success exactly. If the strategy succeeds, then it means that the suitor with the highest dowry is the k^{th} suitor for some value $k > 50$, and that suitors 51 through $k-1$ propose smaller dowries than the largest of the first 50. What is the likelihood that this occurs for a particular value of k? The likelihood that the k^{th} suitor proposes the highest dowry is 1/100, while the likelihood that the largest dowry amongst the first $k-1$ suitors occurs amongst the first 50 is $50/(k-1)$. Therefore, the total likelihood for this particular value of k is $(1/100) \cdot 50/(k-1)$. Now the likelihood that the strategy succeeds is equal to the sum of the likelihoods for these individual values of k, from $k = 51$ to $k = 100$:

$$\text{Probability of success} = \sum_{k=51}^{100} \frac{1}{100} \cdot \frac{50}{k-1} = \frac{50}{100} \sum_{k=51}^{100} \frac{1}{k-1}.$$

How do we estimate this quantity? Note that $\sum_{k=51}^{100} \frac{1}{k-1}$ is a Riemann sum that estimates (from above) the value of the integral $\int_{50}^{100} dx/x = \ln(100) - \ln(50) = \ln(100/50)$. In fact, this approximation is good to within 1/100, noting that $\sum_{k=51}^{100} \frac{1}{k-1} - \frac{1}{50} + \frac{1}{100} = \sum_{k=51}^{100} \frac{1}{k}$ estimates the value of this integral from below. It follows that $P \approx (50/100) \cdot \ln(100/50) = \frac{1}{2}\ln(2) \approx 0.347$. So the likelihood of success is at least 34.7%, or more than one in three!

That worked so well, we are curious about varying our strategy by letting the first k suitors pass and then choosing the first suitor to follow whose dowry exceeds that of the first k. We've seen that $k = 50$ ensures a fairly high likelihood of success. Which value of k should we choose in order to maximize this likelihood?

The same calculation as before shows that the probability of selecting the suitor with the highest dowry is very nearly $k/n \cdot \ln(n/k)$, where $n = 100$. Following the guideline that it is often easier to optimize continuous functions than it is discrete ones, we set $x = \frac{n}{k}$ and seek to maximize the function $f(x) = x^{-1}\ln(x)$ for $x \geq 1$. We calculate the derivative using the product rule: $f'(x) = -x^{-2}\ln(x) + x^{-2} =$

$x^{-2}(\ln(x)-1)$. This vanishes at $x=e$, so e is a critical point of f. We infer that it is a global maximum on $[1,\infty)$ by noting that e is the unique critical point on this interval, $f(e)>0$, and $f(1)=\lim_{x\to\infty} f(x)=0$ (no need for second derivatives!).

In summary, we want to choose $k \approx n/e \approx 0.368n$. That is, we let the first 36.8% of the suitors pass and then select the next one with a higher dowry than any that have already been proposed. For $n=100$, this means the first 37. The likelihood of success is then, somewhat coincidentally, about $e^{-1}\ln(e) \approx 36.8\%$!

The dowry problem has many reincarnations, both frivolous and serious, in life around us. Here are a few of them.

• *The secretary problem.* Suppose that a company is interviewing candidates for a position. It has 100 applicants and wants to select the best one. The company must accept the possibility that if it interviews a candidate and does not make him or her an offer immediately, then it may lose that person for good. Similarly, once it makes a candidate an offer, it is unable to consider the subsequent candidates. One potential strategy for the company to take is to interview the first 37 candidates, get a sense for the field, and then offer the position to the best candidate they interview thereafter.

• *Cheap gasoline.* Suppose that you are driving (updated for a modern audience: and your smartphone is not charged) and in need of gasoline. You pull off of the highway and are willing to drive upwards of ten minutes in search of gas before turning back to the highway. You pass a succession of gas stations and have to decide at which one to stop in order to get the best value. Then you should drive ten/e minutes = tn minutes ≈ 3:40, make a note of the lowest price p, and then stop at the first gas station with a lower price than p.

• *Finding a soulmate.* Most importantly, perhaps, suppose that you are now about 20 years old and that you would like to choose a spouse by the age of 30. (After that, you are committed to a lifetime of solitude, as potential mates will have been paired off already and your friends will be absorbed into their own families.) What you should do is to start dating immediately: the more people, the better. Get a good sense for how much you like each person, but do not waste too much time on anyone for the first 3.68 years; just date them long enough to assess how much you like each one. Then, after 3.68 years (June 2017, figuring from the date of this talk), get serious, and propose to the first person whom you like more than the best of the first batch. The likelihood you'll choose the best one of the people you were likely to date is about 37%! Of course, if you're smart, you'll keep returning the phone calls of your favorites from the first bunch, just in case your initial strategy falls through. Let me know how it goes!

Postscript. The dowry problem is well-known, and you can easily locate various incarnations of it (like the secretary problem) online. Wikipedia will tell you that it may be due to Merrill Flood from 1949, and that it was first popularized in print by Martin Gardner in 1960; it may even go as far back as to Arthur Cayley or Johannes Kepler! Wikipedia will also point you towards some serious generalizations and applications of the dowry problem. I first learned about it in the above form from David C. Kelly in his talk "Sex and the single statistician" at the Hampshire College Summer Studies in Mathematics.

– Josh Greene