Probabilistic Approach to Cinderella Problem

Charles Jeon
University of Pennsylvania
May 5, 2012

Abstract

This short paper serves as a probabilistic approach to one of the classical problems in marriage. While this approach is more generally referred as the Secretary Problem, we apply it in woman’s point of view in accepting proposals with additional assumptions.

Consider a pool of \( n \) men proposing to a girl in order. Assuming that every girl has her own Cinderella story, we can assume, without loss of generality, that there exists a Prince that she is waiting for. Furthermore, we assume that the girl only has one chance to accept the proposal for simplicity. This implies that if she declines, she will never be able to turn back on her decision and if she does accept the proposal, she would not be able to meet the other men. Therefore, in the girl’s point of view, her strategy is to maximize her chances of selecting her Prince from a pool of \( n \) men by selecting the best offer.

With the assumptions aside, let us denote event \( \mathcal{P} \) as the event that the girl successfully accepts the proposal from the Prince and events \( \{ \mathcal{A}_i : 1 \leq i \leq n \} \) as the events that the Prince is in the \( i \)-th order. Clearly, the events \( \{ \mathcal{A}_i \} \) are disjoint. Therefore as \( \mathcal{P} \subseteq \{ \mathcal{A}_i \} \), with elementary Probability Theory, the probability of the girl selecting the Prince is,

\[
\Pr \{ \mathcal{P} \} = \Pr \{ \mathcal{P} | \mathcal{A}_1 \} \Pr \{ \mathcal{A}_1 \} + \cdots + \Pr \{ \mathcal{P} | \mathcal{A}_n \} \Pr \{ \mathcal{A}_n \}
\]

\[
= \sum_{i=1}^{n} \Pr \{ \mathcal{P} | \mathcal{A}_i \} \Pr \{ \mathcal{A}_i \} \tag{1}
\]

Now, suppose that the girl rejects first \( k \)-th proposals and accepts any proposals better than her past \( k \) offers from the \( n-k \) remaining proposals. This implies that

\[
\Pr \{ \mathcal{P} | \mathcal{A}_1 \} = \cdots = \Pr \{ \mathcal{P} | \mathcal{A}_k \} = 0
\]

\[\text{(2)}\]

**Proposition 1.** If the girl’s Prince was in her first \( k \)-th proposal, her strategy fails.

Since the girl rejected her first \( k \)-th offers, Prince’s included, there does not exist any better proposals for the remaining offers, her strategy fails.

**Proposition 2.** For \( k < r \leq n \), \( \Pr \{ \mathcal{P} | \mathcal{A}_r \} = \frac{k}{r-1} \)

This can be explained by a simple counting argument. Consider the cases below.

**Case 1.** Assume \( r = k + 1 \). The Prince proposes immediately after the girl rejects the first \( k \) offers. The girl will accept the proposal and meets her goal. Therefore, we have \( \Pr \{ \mathcal{P} | \mathcal{A}_{k+1} \} = \frac{k}{k} = 1 \).

**Case 2.** Assume \( k + 1 < r \leq n \). If the Prince is in \( r \)-th position, and there exists any better proposals before him that are better than the first \( k \) offers, the girl will accept an offer not from the Prince and her strategy again, fails. Therefore, if the second best man in the \( r - 1 \) proposals propose to the girl her \( k \) rejections, the girl will accept her proposal from the Prince. Therefore, \( \Pr \{ \mathcal{P} | \mathcal{A}_r \} = \frac{k}{r-1} \).
Assuming that the Prince is distributed uniformly, that is, \( \Pr \{ A_i \} = 1/n \) for all \( 1 \leq i \leq n \), combining Equations 1 and 2 with the Propositions lead us to

\[
\Pr \{ \mathcal{P} \} = \sum_{i=1}^{n} \Pr \{ \mathcal{P} | A_i \} \Pr \{ A_i \} = \frac{k}{n} \sum_{r=k+1}^{n} \frac{1}{r-1}
\]

\[
= \frac{k}{n} \sum_{i=k}^{n-1} \frac{1}{i}
\]

(3)

Since this is a probability measure and \( \Pr \{ \mathcal{P} \} \) is not a monotone sequence, there exists some \( k' \) such that \( \Pr \{ \mathcal{P} \} \) is maximum. For small \( n \), simple numerical computations can be used to solve for \( k' \). However, our interest is the approximate \( k' \) for large \( n \), that is,

\[
k' = \arg \min_k \lim_{n \to \infty} \Pr \{ \mathcal{P} \}
\]

At this point, while the equation above does not seem to make progress, intuitive substitutions with \( x \) as the limit of \( k/n \), and \( t = i/n \) with \( dt = 1/n \) transforms the above to a Riemann approximation to integral,

\[
k' = \arg \min_k \lim_{n \to \infty} \frac{k}{n} \sum_{i=k}^{n-1} \left( \frac{n}{i} \right) \left( \frac{1}{n} \right)
\]

\[
\rightarrow \arg \min_x x \int_x^1 \frac{1}{t} \, dt = -x \log_e x
\]

(4)

Elementary calculus for Equation 4 gives \( x = 1/e \simeq 0.3678 \) which for finite \( n \), the optimal \( k' = n/e \simeq 0.3678n \). With \( n = 100 \), this translates to \( k' \simeq 37 \). Therefore, if 100 men propose to a girl, her best possible chance of action is to reject the first 37 proposals. Since 100 proposals are high unlikely for an average girl to receive, we consider \( n = 5 \), which would be a safe number that on average girl receives proposals at max. In this case, she might be able to reject for the first 1, but if she would most likely accept the second proposal if it had been better than the first one.

References