Scientific Research

Probability Brainteasers for College Students

Sahand Rabbani[1], Hossein Shahmohamad[2]

[1]Independent, Rochester, USA
[2]School of Mathematical Sciences, Rochester Institute of Technology, Rochester, USA
Email: hxssma@rit.edu

Abstract

We present eleven interesting problems in probability along with their solutions. These problems can be used as challenging and entertaining questions for high school or first-year college students.

Keywords

Probability, Statistics

1. Introduction

Quantitative brainteasers are more than just fascinating. They are parts of the natural pastime in engineering and science. The exact solution to a few may require greater computing power, but a satisfactory response without computer aid is provided in this paper. Two excellent references for material leading to these problems can be found in [1] or [2].

1.1. Triangle

A stick of length L is randomly cut in two places. What is the probability that the three resulting pieces can form a triangle?

Solution: By depicting the event space as an L-by-L square in (x, y)-space, where the point (X, Y) represents cuts at distances X and Y along the stick, it becomes evident that the probability is equal to the portion of the area bounded by certain linear constraints that ensure a triangle can be formed. This probability is $\frac{1}{4}$.

1.2. Disorganized Meet-Up

Two people have agreed to meet at a particular location between the hours of noon and 1 PM. Each individual will arrive uniformly and independently at a random time in this interval. If an individual arrives and the other is not there, the arriving party will wait for 20 minutes. If the other does not arrive within these 20 minutes, the waiting parting will leave. What is the probability that the two people will meet?

Solution: Plotting the event space on a 1 hr-by-1 hr grid and applying constraints that bound the area in which the two people meet reveals that the probability is $\frac{5}{9}$.

1.3. Collect Them All

Consider a brand of cereal that advertises a special prize in each of its cereal boxes. Suppose that there are n possible prizes and that any box of cereal contains each of these n prizes with equal probability, but which prize it contains is unknown until the box is opened. What is the expected number of cereal boxes that one must open to obtain all n prizes?

Solution: Define the random variable X_n as the number of trials, or cereal box openings, until all n prizes are obtained. We seek $E[X_n]$. Consider the random variables Y_i for $i = 1, 2, \cdots, n$, where Y_i only has significance after n - i prizes have been obtained. The random variable Y_i is the number of trials required to obtain one of the i remaining prizes. We can thus write X_n as

$$X_n = Y_n + Y_{n-1} + \cdots + Y_1 = \sum_{i=1}^{n} Y_i$$

By the linearity of expectation, we have

$$E[X_n] = E[Y_n] + E[Y_{n-1}] + \cdots + E[Y_1] = \sum_{i=1}^{n} E[Yi]$$

We note that each Y_i is a geometric random variable where the probability of success p_i is $\frac{i}{n}$. The simplest case is $Y_n = 1$. When none of the prizes have been obtained, the expected number of trials to obtain one of the n prizes is deterministic: open one cereal box and you have one of these n prizes. In general, the expected value of a geometric random variable with probability of success p is $\frac{1}{p}$. Thus, the expected value of X_n is

$$E[Xn] = \sum_{i=1}^{n} E[Yi] = \sum_{i=1}^{n} \frac{1}{\frac{i}{n}} = n \sum_{i=1}^{n} \frac{1}{i}$$

1.4. The Cuckoo Train

There are n passengers who have booked a seat on a full-capacity train ride. They are lined up to board the train, each passenger with a designated seat. Without loss of generality, we say that the nth passenger in line is assigned the nth seat. The first passenger to board, however, is an irrational man with no regard for Kants categorical imperative. Instead of sitting at his assigned seat, he randomly picks among the n seats. Each subsequent passenger to board, then, first looks to see if his assigned seat is available. If it is, he sits there. Otherwise, he picks randomly from the remaining seats. What is the probability that the last passenger will sit in his assigned seat?

Solution: Let us call the probability that the final passenger sits in his assigned seat p. Also, let us define the useful quantity p_i, which is the probability that the last person will sit in his assigned seat given an i-person subtrain with the initial person picking a random seat. Clearly, $p = p_n$. Now, consider that when the first person boards, there is a probability $\frac{1}{n}$ that he will pick his own seat, resulting in proper boarding with the final

passenger sitting in his assigned seat as well. With probability $\frac{1}{n}$, he will pick the second person's seat, after which the second passenger sees that his seat is taken and must then pick a random seat. In general, with the same probability $\frac{1}{n}$, the first passenger will pick the i th seat, after which boarding will proceed normally until the i th passenger finds that his seat is taken, and he must pick a random seat:

$$p_n = \frac{1}{n}\left(1 + \sum_{i=2}^{n-1} p_i\right)$$

Thus,

$$np_n = 1 + \sum_{i=2}^{n-1} P_i = p_{n-1} + \left(1 + \sum_{i=2}^{n-1} p_i\right)$$

Also, $(n-1)p_{n-1} = 1 + \sum_{i=2}^{n-1} p_i$ which implies $p_n = p_{n-1} = \cdots = p_2$.

Clearly, $p_2 = \frac{1}{2}$. This is the case where there are two seats and the first person randomly picks between his own seat and the seat of the last (second) passenger. Thus, $p = p_n = \frac{1}{2}$. The final passenger will sit in his assigned seat with probability $\frac{1}{2}$. Another solution considers that the only relevant event is whether a passenger sits in the first man's seat, in which case the last man will get his seat, or sits in the last man's seat, in which case the last man will not get his seat. If a passenger sits in any other seat, the same consideration is then deferred to the next boarding passenger. Since the probability of a man sitting in either the first man's seat or the last man's seat is equal, the probability of the last man getting his seat is equal to the probability of the last man not getting his seat, *i.e.*, $\frac{1}{2}$.

1.5. Heads, I Win

There is a two-headed coin and fifteen fair coins. One coin is chosen at random and flipped, coming up heads. What is the probability that this coin is the two-headed coin?

Solution: This is a simple application of Bayes rule: there are 17 head faces, two of which belong to the two-headed coin. The conditional probability of a head face belonging to the two-headed coin is $\frac{2}{17}$.

1.6. Tails, You Lose

1) There is a loaded coin that flips tails with probability p, where p is known only to be distributed uniformly on the interval [0,1]. The coin is flipped five times, each time turning up tails. What is the best estimate of p?

Solution: This is another application of Bayes rule:

$$E\left[p|T^5\right] = \int_0^1 \Pr\left[p|T^5\right] p\,\mathrm{d}p$$

$$= \int_0^1 \frac{\Pr\left[T^5|p\right]\Pr[p]}{\int_0^1 \Pr\left[T^5|p\right]\Pr[p]\mathrm{d}p} p\,\mathrm{d}p$$

$$= \frac{\int_0^1 p^6\,\mathrm{d}p}{\int_0^1 p^5\,\mathrm{d}p} = \frac{6}{7}$$

2) In general, what is the best estimate of p conditional on flipping the coin n times and observing exactly m tails?

Solution: The general solution is

$$E\left[p\big|T^m H^{n-m}\right] = \int_0^1 \Pr\left[T^m H^{n-m}\right] p\,\mathrm{d}p$$

$$= \int_0^1 \frac{\Pr\left[T^m H^{n-m}\big|p\right]\Pr\left[p\right]}{\int_0^1 \Pr\left[T^m H^{n-m}\big|p\right]\Pr\left[p\right]\mathrm{d}p} p\,\mathrm{d}p = \frac{\int_0^1 \binom{n}{m} p^{m+1}(1-p)^{n-m}\,\mathrm{d}p}{\int_0^1 \binom{n}{m} p^m (1-p)^{n-m}\,\mathrm{d}p}$$

$$= \frac{\int_0^1 \binom{n}{m}\sum_{k=0}^{n-m}\binom{n-m}{k}(-1)^k p^{k+m+1}\mathrm{d}p}{\int_0^1 \binom{n}{m}\sum_{k=0}^{n-m}\binom{n-m}{k}(-1)^k p^{k+m}\mathrm{d}p} = \frac{\sum_{k=0}^{n-m}\binom{n-m}{k}\dfrac{(-1)^k}{k+m+2}}{\sum_{k=0}^{n-m}\binom{n-m}{k}\dfrac{(-1)^k}{k+m+1}}$$

1.7. Metropolitan Transit Authority

There are N sequentially numbered bus routes in a city. Knowing only that N is uniformly distributed on $[a,b]$, one randomly observes a bus route numbered $X \le b$. What is the best estimate of N based on this observation?

Solution: Let $f(X|n)$ be the distribution of X conditional on $N = n$ and let $g(n)$ be the distribution of N. Assume that X and N are continuous random variables for simplicity. The conditional expectation of N given X is

$$E\left[N\big|X\right] = \frac{\int_{-\infty}^{\infty} f(X|n) g(n) n\,\mathrm{d}n}{\int_{-\infty}^{\infty} f(X|n) g(n)\,\mathrm{d}n}$$

$$= \frac{\int_{-\infty}^{\infty} \dfrac{1}{n} I(X \le n) \dfrac{1}{b-a} I(n \in [a,b]) n\,\mathrm{d}n}{\int_{-\infty}^{\infty} \dfrac{1}{n} I(X \le n) \dfrac{1}{b-a} I(n \in [a,b])\,\mathrm{d}n} = \frac{\int_{\max}^{b} (X,a)\,\mathrm{d}n}{\int_{\max}^{b} (X,a)\,\mathrm{d}n/n} = \frac{B - \max(X,a)}{\ln\left(B/\max(X,a)\right)}$$

1.8. Bulls Eyes

There is an unusual dart board that has two equally sized non-overlapping circles of radius r whose centers are separated by some distance $d > 2r$. In this dart game, if a dart lands inside of any of the circles, one point is scored; otherwise, no points are scored.

1) Where should you aim? Assume that the distribution around your aim is a zero-mean circular Gaussian with variance σ^2.

Solution: Roughly, for small σ, you should aim at the center of one of the circles, and for large σ, you should aim between the two circles.

2) Having determined that the optimal aim point is a function of σ, is this function discontinuous such that there exists some σ above which you should aim between the circles and below which you should aim at one of the centers?

Solution: Based on numerical solutions for the one-dimensional case, the optimal aim is a continuous function of σ but may be non-differentiable at the point of smallest σ for which the optimal aim point is the center of one of the circles.

1.9. Synthetic Bet

There is a seven-game series sports event between Team A and Team B where the series winner is the first team to win four games. Single bets exist for each individual game with fair payout $n:1$ for Team A to win, but there is no bet available for the outcome of the entire series. Considering that bets for each individual game can be entered at any time prior to that game and can be placed in any amount, is there a betting strategy that pays

fair-odds on the outcome of the entire series, and what is that strategy?

Solution: The possible paths and outcomes of this series can be depicted with a binomial tree. The desired payouts (or cash position) at the leaves of the tree are determined from the odds of Team A winning the entire series. The necessary cash position x and the quantity y of the required bet at any intermediate node can be determined by the required cash position of the two possible succeeding nodes. Let a be the required cash position if Team A wins the game at the current node and b be the required cash position if Team B wins the game at the current node. These cash positions can be achieved exactly if the cash position at the current node is

$x = \dfrac{(a+nb)}{(n+1)}$ and a bet for Team A to win is placed in the amount of $y = \dfrac{(a-b)}{(n+1)}$. Using these formulas, the

required cash position and required betting amount can be populated for every node in the binomial tree. Arguably the more challenging part of the problem is to determine the fair odds for Team A to win the entire series. In general, these odds are not n:1. The favored team wins the series with a higher probability than it wins an individual game. This problem is in some ways analogous to the discrete-time binomial pricing of derivatives contracts. At each stage, a replicating portfolio of cash and a risky instrument (a bet on the outcome of the next game) is maintained and subject to dynamic re-balance.

1.10. A Good Deal

A standard deck of 52 cards contains 26 black cards and 26 red cards. You have the opportunity to play a game wherein you may flip over cards from the randomized deck one at a time. At any point, you may stop, and you receive $1 for every black card and pay $1 for every red card that has been turned over. What strategy achieves the maximum expected profit for this game and what is that profit?

Solution: Let $V[i,j]$ represent the expected profit from the optimal strategy in a game where the deck contains i black cards and j red cards. This function can be defined recursively. Under the optimal strategy, the expected profit of the (i,j) game is

$$V[i,j] = \frac{i}{i+j}\big(V[i-1,j]+1\big) + \frac{j}{i+j}\big(V[i,j-1]-1\big)$$

In the degenerate case where there are i black cards and no red cards, the value of the game is clearly $V[i,0]=i$, since the optimal strategy is to turn over all the cards. In the converse where there are j red cards and no black cards, the optimal strategy is not to turn over any cards: $V[0,j]=0$. Given the boundary conditions and the recursive formula, the entire grid of V can be easily populated to establish that the expected profit from the optimal strategy for a standard 52-card deck is approximately $2.6245. The optimal strategy is to flip over cards as long as $V[i,j]\neq 0$ when in the (i,j) game.

1.11. Adverse Selection

You have the opportunity to submit a single bid $b \in [0,1]$ on an object. You know that its intrinsic value v is uniformly distributed on $[0,1]$. If $b \geq v$, you win the object and pay b. If $b < v$, you do not win the object and pay nothing. If you obtain the object, you can perform a costless modification that multiplies its value by $k > 1$ and then proceed to sell the object for kv. What bidding strategy as a function of k maximizes expected profits?

Solution: The conditional expected value of v given that b is a winning bid is $\dfrac{b}{2}$. The conditional profit given that b is a winning bid is then $k(b/2)-b=b(k/2-1)$. The probability that b is a winning bid is b, so the expected profit of a bid b is $b2(k/2-1)$. For $k<2$, this number is non-positive, so the optimal bid is 0. For $k>2$, this number is positive and maximized for the highest bid $b=1$. For $k=2$, the expected profit is always 0, so any bid will do.

References

[1] Evans, J.R. (2007) Statistics, Data Analysis, & Decision Modeling. Pearson/Prentice Hall, Upper Saddle River.

[2] Christensen, H.B. (1992) Introduction to Statistics. Saunders/HBJ, Orlando.

Scientific Research Publishing (SCIRP) is one of the largest Open Access journal publishers. It is currently publishing more than 200 open access, online, peer-reviewed journals covering a wide range of academic disciplines. SCIRP serves the worldwide academic communities and contributes to the progress and application of science with its publication.

Other selected journals from SCIRP are listed as below. Submit your manuscript to us via either submit@scirp.org or Online Submission Portal.