Probability and Statistics

William F. Barnes

April 3, 2020

Contents

Introduction

Probability theory is a branch of mathematics for studying systems with inherent randomness or uncertainty. It works closely with *statistics*, another branch of mathematics concerned with the organization and interpretation of data, along with *combinatorics*, a formal method of counting.

Systems that exhibit random or patternless behavior contain a *stochastic* component. Typical stochastic processes may include flipping a coin, drawing a card from a deck, rolling a dice, or playing darts while blindfolded. A *stochastic event* is any data generated by a stochastic process, and the set of all possible stochastic events is called the system's *sample space.*

1 Events and Probability

1.1 Elementary vs. Compound Events

Events that are considered *elementary* carry one 'unit' of information, loosely speaking. A coin landing on 'heads', or a dice landing on 6 qualify as elementary. However, drawing a Queen of Hearts from a deck of cards carries two units information, and may be interpreted in several ways: 'draw a Queen AND a heart', or 'draw a Queen OR a Heart', or perhaps 'draw NOT a Diamond'. Such events are classified as *compound*.

Borrowing the familiar symbols from elementary logic, we denote the word 'AND' with the 'cap' symbol $\cap$, equivalent to multiplication ($\cdot$). Meanwhile, the word 'OR' uses the 'cup' symbol $\cup$, or sometimes just a plus sign ($+$). The 'NOT' operator is abbreviated by a dash above the symbol, as in 'NOT' $A = \bar{A}$. Any event that is infinitely improbable, impossible, or undefined is denoted by the 'Empty set' symbol, $\emptyset$. In summary:

$$\begin{aligned} A \text{ AND } B &= A \cdot B = A \cap B \\ A \text{ OR } B &= A + B = A \cup B \\ \text{NOT } A &= \bar{A} \\ \text{Empty set} &= \emptyset \end{aligned}$$

The logic of probabilistic analysis is the same as 'ordinary' logic. For instance, the philosophical axiom 'nothing can be and not be simultaneously' is contained in the statement:

$$A \cap \bar{A} = \emptyset$$

1.2 Statistical vs. Classical Probability

A stochastic process that iterates over a very large or infinite number of trials will yield data points across its entire sample space. For all events of type A, the ratio of occurrences N_A over all N events is called the *statistical probability* of event A, defined as

$$P(A) = \lim_{N \to \infty} \frac{N_A}{N} , \tag{1}$$

which strictly has values between 0 and 1, inclusive. All other events B, C, etc., are represented by the symbol $\bar{A}$ ('NOT' A), and obey:

$$N_A + N_{\bar{A}} = N \qquad\qquad P\left(\bar{A}\right) = 1 - P(A)$$

A definition of that skirts around the invocation of $N \to \infty$ is called the *classical probability*. By exploiting the symmetry of a balanced six-sided dice, we reason that the probability of landing on any given number is precisely 1/6. Similarly for a deck of 52 cards, the probability of drawing the Queen of Hearts is 1/52.

Problem 1

A bank account password has format ABCD, where each letter represents any integer from 0 to 9, inclusive. What is the probability of randomly guessing the password?

Solution 1

$$N = 10000 \qquad \rightarrow \qquad P\left(N_A\right) = \frac{1}{N} = \frac{1}{10000}$$

Problem 2

A bank account password has format ABCD, where each letter represents any integer from 0 to 3, inclusive. What is the probability of randomly guessing the password?

Solution 2

All two-digit arrangements solved by AB are contained in

$$\omega = (00, 01, 02, 03, 10, 11, 12, 13, 20, 21, 22, 23, 30, 31, 32, 33) \ ,$$

where we observe that all four-digit arrangements are contained on an $\omega \times \omega$ grid having total arrangements $N = 16^2 = 256$, or

$$P\left(N_A\right) = \frac{1}{N} = \frac{1}{256} \ .$$

1.3 Mutually Exclusive Events

A pair of *mutually exclusive events* A and B are those that cannot occur simultaneously. Their coincidence can only belong to the *empty set* as

$$A \cap B = \emptyset \ .$$

If two events are mutually exclusive, the probability of either event occurring is

$$P\left(A \text{ or } B\right) = P\left(A \cup B\right) = P\left(A\right) + P\left(B\right) \ . \tag{2}$$

Non-mutually exclusive events are those that cause 'double counting' in $P\left(A \cup B\right)$, and are adjusted by subtracting the joint probability:

$$P\left(A \text{ or } B\right) = P\left(A \cup B\right) = P\left(A\right) + P\left(B\right) - P\left(A \cap B\right) \tag{3}$$

Problem 3

Use equation (2) to calculate the probability of rolling a 3 or a 4 on a six-sided dice.

Solution 3

$$P = \frac{1}{6} + \frac{1}{6} = \frac{2}{3}$$

Problem 4

Calculate the probability that a random three-card hand drawn from a 52-card deck contains the Queen of Hearts.

Solution 4

$$P = \frac{1}{52} + \frac{1}{52} + \frac{1}{52} = \frac{3}{52}$$

Problem 5

From a 52-card deck, use equation (3) to calculate the probability of drawing a Heart or a Face card, or one that is both.

Solution 5

$$P = \frac{13}{52} + \frac{12}{52} - \frac{3}{52} = \frac{22}{52}$$

1.4 Independent Events

Consider two events A and B, having respective probabilities $P(A)$, $P(B)$ of occurring. Taking the events as simultaneous, there are two trivial cases for the probability $P(A \cap B)$:

$$\begin{aligned} P(A \cap B) &= 0 && \text{if} && A \cap B = \emptyset \\ P(A \cap B) &= P(A) = P(B) && \text{if} && A = B \end{aligned}$$

Proceeding with A and B as *independent* events, we begin with the definition (1) of statistical probability for the compound event

$$P(A \cap B) = \lim_{N \to \infty} \frac{N_{A \cap B}}{N},$$

where in the $N \to \infty$, limit the quantity $N_{A \cap B}$ becomes $N_A \cdot P(B)$. We deduce that the probabilities multiply as

$$P(A \cap B) = P(A) \cdot P(B) \tag{4}$$

for independent events.

Problem 6

Calculate the probability of two fair coin tosses each landing on ‘tails’.

Solution 6

$$P(T \cap T) = P(T) \cdot P(T) = (1/2)(1/2) = 1/4$$

Problem 7

(i) From a 52-card deck, what is the probability $P(Q)$ of randomly drawing a Queen? (ii) From a 52-card deck, what is the probability $P(H)$ of randomly drawing a Heart? (iii) Calculate $P(Q \cap H)$ using equation (4).

Solution 7

$$P(Q) = \frac{1}{4} \qquad\qquad P(H) = \frac{1}{13}$$
$$P(Q \cap H) = P(Q) \cdot P(H) = \frac{1}{4} \cdot \frac{1}{13} = \frac{1}{52}$$

Problem 8

Calculate the probability of rolling a 3 or greater m times in a row on a six-sided dice. Hint: Calculate the probability of generating m consecutive rolls of NOT greater than three, and then exploit $P(A) = 1 - P(\bar{A})$.

Solution 8

$$P(A_1) = \frac{2}{3} \qquad\qquad P(\bar{A}_1) = 1 - \frac{2}{3} = \frac{1}{3}$$
$$P(\bar{A}_1 \cdots \bar{A}_m) = P(\bar{A}_1) \cdots P(\bar{A}_m) = \left(\frac{1}{3}\right)^m$$
$$P(A_1 \cdots A_m) = 1 - P(\bar{A}_1 \cdots \bar{A}_m) = 1 - \left(\frac{1}{3}\right)^m$$

1.5 Conditional Probability

The probability for event B to occur given the occurrence of event A is called the *conditional probability*, denoted $P(B|A)$, enunciated as 'B given A'. Using definition (1), the event B occurring given condition A is

$$P(B|A) = \lim_{N \to \infty} \frac{N_{A \cap B}}{N_A}\,.$$

To deal with the $N_{A \cap B}$ term, write a second equation containing it, namely

$$P(A \cap B) = \lim_{N \to \infty} \frac{N_{A \cap B}}{N}\,.$$

Divide the two equations simplify:

$$P(A \cap B) = P(B|A)\,P(A) \tag{5}$$

Note that if events A and B are independent, equation (5) reduces to equation (4) again.

Problem 9

Using conditional probabilities, calculate the chance that a random three-card hand drawn from a 52-card deck contains the Queen of Hearts. Hint: Define the event $\bar{A}$ of NOT drawing the Queen of Hearts, and exploit $P(A) = 1 - P(\bar{A})$.

Solution 9

$$\bar{A} = \bar{A}_1 \cdot \bar{A}_2 \cdot \bar{A}_3$$

$$P\left(\bar{A}_1 \cap \bar{A}_2\right) = P\left(\bar{A}_1\right) P\left(\bar{A}_2 | \bar{A}_1\right) = \frac{51}{52} \cdot \frac{50}{51} = \frac{50}{52}$$

$$P\left(\bar{A}_1 \cap \bar{A}_2 \cap \bar{A}_3\right) = P\left(\bar{A}_1 \cap \bar{A}_2\right) P\left(\bar{A}_3 | \bar{A}_1 \cap \bar{A}_2\right) = \frac{50}{52} \cdot \frac{49}{50} = \frac{49}{52}$$

$$P(A) = 1 - P\left(\bar{A}\right) = 1 - \frac{49}{52} = \frac{3}{52}$$

Problem 10

In a 52-card deck, calculate the probability that the first three cards are Kings. (This is a repeat of an earlier problem. Attain a solution using conditional probabilities.)

Solution 10

$$P = P(K) P(K|K) P(K|(K|K)) = \frac{4}{52} \cdot \frac{3}{51} \cdot \frac{2}{50} \approx 0.000181$$

Problem 11

In a 52-card deck, calculate the probability that the first three cards are KQJ, in that order, with mixed suits allowed. (This is a repeat of an earlier problem. Attain a solution using conditional probabilities.)

Solution 11

$$P = P(K) P(Q|K) P(J|(Q|K)) = \frac{4}{52} \cdot \frac{4}{51} \cdot \frac{4}{50} \approx 0.000483$$

Problem 12

Suppose a pair of six-sided dice are rolled, landing on faces X and Y, respectively. What is the probability that $X = 2$ given that $X + Y \leq 5$?

Solution 12

Denote condition B as $X + Y \leq 5$, and event A as $X = 2$. There is a $1/6$ chance of rolling $X = 2$, and meanwhile $10/36$ of all outcomes satisfy $X + Y \leq 5$. The number of configurations satisfying *both* can be readily named as $(2, 1)$, $(2, 2)$, $(2, 3)$, a total of three outcomes, indicating $P(A \cap B) = 3/36$. Using the formula for conditional probability, we have

$$P\left(X = 2 \middle| X + Y \leq 5\right) = \frac{3/36}{10/36} = \frac{3}{10} .$$

Radioactive Decay

An unstable atom is one that expels energy by ejecting a subatomic particle or photon. Having no internal time-keeping mechanism, an unstable atom is entirely 'unaware' of its absolute age. Supposing the observation of an unstable atom begins at $t = 0$, the conditional probability of the atom decaying in a small time window Δt after time $t > 0$ is

$$P^{\text{decay}}_{\Delta t/t} = \tau^{-1} \Delta t ,$$

where τ^{-1} is a constant related to (but not precisely equal to) the statistical half-life of the element, defined such that $\Delta t \ll \tau$.

The probability of the atom being 'still alive' in the interval Δt is

$$P_{\Delta t/t}^{\text{alive}} = 1 - \Delta t/\tau \ .$$

Decompose the entire 'alive' state into a product of conditional probabilities by slicing the time t into n identical copies of the short interval Δt as

$$P^{\text{alive}}\left(t\right) = P_{\Delta t/t_1}^{\text{alive}} \cdot P_{\Delta t/t_2}^{\text{alive}} \cdots P_{\Delta t/t_n}^{\text{alive}} = \left(1 - \frac{t}{n\tau}\right)^n \ .$$

Letting $n \to \infty$ permits use of the identity

$$\lim_{n\to\infty} \left(1 + \frac{A}{n}\right)^n = e^A \ .$$

It follows that the probability that a single unstable atom will still be 'alive' obeys

$$P\left(t\right) = e^{-t/\tau} \ .$$

Missing Face

Problem 13

Consider a six-sided dice modified to refuse landing with the 2 facing up. When attempting to land on 2, the dice self-tosses to roll again. Prove that the statistical probability of landing on any valid number 1, 3, 4, 5, 6 converges to 1/5.

Solution 13

Denote B as the event 2, and denote A as any valid event 1, 3, 4, 5, 6. With a single roll, the probabilities of A or B occurring are easy to write down:

$$p_1\left(A\right) = \frac{1}{6} \qquad\qquad p_1\left(B\right) = \frac{1}{6}$$

Of course event B is unstable and induces a re-roll, which, has a 1/6 chance of generating event A again, and the same chance for event B:

$$p_2\left(A|B\right) = \frac{1}{6} \cdot \frac{1}{6} \qquad\qquad p_2\left(B|B\right) = \frac{1}{6} \cdot \frac{1}{6}$$

With event B comes another re-roll, and we stack on the conditional probabilities as

$$p_3\left(A|B|B\right) = \frac{1}{6} \cdot \frac{1}{6} \cdot \frac{1}{6} \qquad\qquad p_3\left(B|B|B\right) = \frac{1}{6} \cdot \frac{1}{6} \cdot \frac{1}{6} \ ,$$

and the pattern is clear.

It follows that A may not occur on a single roll, but could potentially occur after an infinite string of B-events as

$$P\left(A\right) = p_1\left(A\right) + p_2\left(A|B\right) + p_3\left(A|B|B\right) + p_4\left(A|B|B|B\right) + \cdots \ ,$$

which simplifies to

$$P(A) = \frac{1}{6} \cdot \left(1 + \left(\frac{1}{6}\right) + \left(\frac{1}{6}\right)^2 + \left(\frac{1}{6}\right)^3 + \cdots \right).$$

In the infinite limit, the series in parentheses converges to 6/5. The probability of event A occurring is therefore:

$$P(A) = \frac{1}{6} \cdot \frac{6}{5} = \frac{1}{5}$$

1.6 Bayes' Theorem

Observe by construction that $P(A \cap B) = P(B \cap A)$, thus we exploit the symmetry in equation (5) to write a second version with all A and B symbols swapped. Eliminate $P(A \cap B) = P(B \cap A)$ to arrive at *Bayes' theorem*:

$$P(B|A) = \frac{P(A|B)\, P(B)}{P(A)} \tag{6}$$

Laptop Repair Shop

Source: CMPSCI240 UMass Amherst 2013

Problem 14

You work in a laptop repair shop. 80% of laptops brought in have been dropped, 15% of laptops have had a drink spilled on them, and 5% of laptops have a variety of other problems. A customer drops off a laptop and doesn't tell you what happened to it. You notice the laptop is emitting a slight coffee-like smell. Based on your knowledge of broken laptops, you estimate that 20% of dropped laptops have a slight coffee-like smell, 65% of laptops that have had something spilled on them have a slight coffee-like smell, and 5% of laptops that have some other problem have a slight coffee-like smell.

Provide labels for the events described in the problem. Find the probability that the laptop had something spilled on it given that is has a slight coffee-like smell.

Problem 15

After closer inspection, you note that the laptop has no cracks on the case. Based on your knowledge of broken laptops, you estimate that 80% of dropped laptops have cracked cases, 11% of laptops that have had something spilled on them have cracked cases, and 9% of laptops that have some other problem have cracked cases.

If the probability that a laptop smells like coffee and the probability that a laptop has a cracked case are conditionally independent of each other given the case of the damage (drop, spill, or other), what is the probability that the laptop had something spilled on it if it has a slight coffee-like smell and no cracks in the case?

Solution 14

Denote D for 'drop', S for 'spill', and O for 'other'. Let the letter F denote 'slight coffee-like smell'. The information provided in the problem may be written:

$$P(D) = .80 \qquad P(S) = .15 \qquad P(O) = .05$$

$$P(F|D) = .20 \qquad P(F|S) = .65 \qquad P(F|O) = .05$$

We further deduce:

$$P(F) = P(F|D)\,P(D) + P(F|S)\,P(S) + P(F|O)\,P(O) = .26$$

To answer the question, we need to compute $P(S|F)$, which is the inversion of $P(F|S)$. Thus:

$$P(S|F) = \frac{P(S)\,P(F|S)}{P(F)} = .375$$

<u>Solution 15</u>

The problem asks us to evaluate $P\left(S|F \cap Z^C\right)$, where Z denotes 'crack' and Z^C denotes 'no crack'. Due to the independence between F and Z^C, the term $P\left(F \cap Z^C|S\right)$ decouples into $P(F|S)\,P\left(Z^C|S\right)$. Using Bayes' theorem we write

$$P\left(S|F \cap Z^C\right) = \frac{P(S)\,P(F|S)\,P\left(Z^C|S\right)}{\sum_{i=D,S,O} P\left(F \cap Z^C|X_i\right)P(X_i)},$$

and meanwhile the problem has stated

$$P(Z|D) = .80 \qquad P(Z|S) = .11 \qquad P(Z|O) = .09\,,$$

or equivalently:

$$P\left(Z^C|D\right) = 1 - .80 \qquad P\left(Z^C|S\right) = 1 - .11 \qquad P\left(Z^C|O\right) = 1 - .09$$

Evaluating $P\left(S|F \cap Z^C\right)$ is now straightforward:

$$P\left(S|F \cap Z^C\right) = \frac{.15 \times .65 \times .89}{.20 \times .20 \times .80 + .65 \times .90 \times .15 + .05 \times .91 \times .05} = .7169$$

2 Combinatorics

2.1 Arrangements

The number A_n of all *arrangements* of n distinguishable (non-repeated) elements is equal to the factorial of the total number of elements:

$$A_n = n! \tag{7}$$

If the set of n elements contains any number m identical members, the number of arrangements over-counts by a factor of m-factorial, which must be divided out:

$$A_n^m = \frac{n!}{m!} \tag{8}$$

Problem 16

Consider the set of twelve elements $ACEFGILMNTUX$. What is the probability that a random arrangement of the elements will spell out $MAGNETICFLUX$?

Solution 16

$$P\left(MAGNETICFLUX\right) = \frac{1}{12!}$$

Problem 17

Consider the word $FLUXELECTRIC$. What is the probability that a random arrangement of the letters will spell out $ELECTRICFLUX$?

Solution 17

$$P\left(ELECTRICFLUX\right) = \frac{2!2!2!}{12!}$$

2.2 Permutations

For a set of width n, partition each of the $n!$ arrangements into two bins such that one bin contains the first m elements in the arrangement, and the other bin contains the remaining $n-m$ elements. For each of the m elements in the first bin, the unused elements in the second bin are subject to $(n-m)!$ arrangements. Dividing out this number yields the *permutation* number:

$$P_n^m = \frac{n!}{(n-m)!} \tag{9}$$

Problem 18

A door keypad is unlocked by a code of four different integers between 0 to 9, inclusive. (The same integer cannot be used twice.) What is the probability of randomly guessing the password?

Solution 18

For a $k = 4$ digit password drawing (and consuming) from $N = 10$ integers, observe that N of them are available for the first digit A. $N - 1$ of the digits are available for the second digit B, and so on, with the k^{th} digit selecting from $N - k + 1$ remaining integers. Thus:

$$\Omega = N\left(N-1\right)\left(N-2\right)...\left(N-k+1\right) = \frac{N!}{(N-k)!} = \frac{10!}{6!} = 5040$$

Problem 19

In a 52-card deck, calculate the probability that the first three cards are Kings.

Solution 19

$$P\left(KKK\right) = \frac{4!/\left(4-3\right)!}{52!/\left(52-3\right)!} = \frac{4\cdot 3\cdot 2}{52\cdot 51\cdot 50} \approx 0.000181$$

Problem 20

In a 52-card deck, calculate the probability that the first three cards are KQJ, in that order, with mixed suits allowed.

Solution 20

$$P(KQJ) = \frac{4^3}{52!/(52-3)!} = \frac{4^3}{52 \cdot 51 \cdot 50} \approx 0.000483$$

Birthday Problem

Problem 21

Consider a classroom of total population N. What is the probability that any two people were born on the same day?

Solution 21

Begin with the trivial case $N = 2$, in where there is a $1/365$ chance of a common birthday:

$$P(2) = \frac{1}{365} = 1 - \frac{364}{365}$$

The result is written in the form $1 - X$ so we may focus on X, the probability of *no* common birthday.

A third person entering the system, making $N = 3$, has $365 - 2 = 363$ available days to avoid a common birthday. The probability becomes

$$P(3) = 1 - \frac{364}{365} \cdot \frac{363}{365},$$

and the pattern is now obvious. For total population N, the probability that some pair of people share a birthday is:

$$P(N) = 1 - \frac{365}{365} \cdot \frac{364}{365} \cdot \frac{363}{365} \cdots \frac{(365 - N + 1)}{365} = 1 - \frac{365!}{365^N (365 - N)!}$$

Note that X has been expressed as a recursion of conditional probabilities

$$X(n|n-1) = \frac{365 - (n-1)}{365} \qquad \rightarrow \qquad X(N) = \prod_{n=2}^{N} X(n|n-1),$$

which could also have been written directly by the permutation formula

$$X(N) = \frac{P_{365}^{N}}{365^N} = \frac{365!}{365^N (365 - N)!}.$$

Following is a list of various populations N with their corresponding $P(N)$:

N	P(N)
5	2.71%
10	11.7%
20	41.1%
23	50.7%
30	70.6%
50	97.0%

Remarkably, the population need only be 23 in order for there to be a 50% chance that any two people share a birthday.

2.3 Combinations

When the precise order of elements in the 'm' bin does not matter, the list of permutations is overpopulated by a factor of $m!$. Dividing out this number, we attain the number of *combinations*:

$$C_n^m = \frac{n!}{m!\,(n-m)!} \tag{10}$$

The numbers C_n^m are also called the *binomial coefficients*.

Problem 22

From a 52-card deck, a five-card hand is drawn at random. How many five-card hands are possible?

Solution 22

$$C_{52}^5 = \frac{52!}{5!\,(52-5)!}$$

Problem 23

From a 52-card deck, calculate the probability of drawing a royal flush (A-K-Q-J-10) in any order in any one suit.

Solution 23

$$P\,(RF) = \frac{4}{C_{52}^5} = \frac{1}{649740} \approx 0.00000154$$

Mega Millions

Source: Durango Bill
`http://www.durangobill.com/MegaMillionsOdds.html`

In a lottery game, the winning numbers are five non-repeating integers between 1 and 75, inclusive, along with one bonus integer between 1 and 15, inclusive. Guessing the five winning numbers at random, let us calculate the probability $P = (n, b)$ that n of the guessed numbers match the winning numbers, with or without the bonus b also being correctly guessed.

As an application of combinatoric analysis, it follows that there are $C_{75}^5 = 17,259,390$ ways to guess the winning five numbers, and $C_{15}^1 = 15$ choices for the bonus number. Following are the probabilities of guessing any number of winning numbers, with and without the bonus.

$$P\,(5,1) = \frac{1}{C_{75}^5 \cdot C_{15}^1} = \frac{1}{258,890,850} \qquad P\,(4,1) = \frac{C_5^4 \cdot C_{70}^1}{C_{75}^5 \cdot C_{15}^1} \approx \frac{1}{739,688}$$

$$P(3,1) = \frac{C_5^3 \cdot C_{70}^2}{C_{75}^5 \cdot C_{15}^1} \approx \frac{1}{10,720}$$

$$P(2,1) = \frac{C_5^2 \cdot C_{70}^3}{C_{75}^5 \cdot C_{15}^1} \approx \frac{1}{473}$$

$$P(1,1) = \frac{C_5^1 \cdot C_{70}^4}{C_{75}^5 \cdot C_{15}^1} \approx \frac{1}{56}$$

$$P(0,1) = \frac{C_5^0 \cdot C_{70}^5}{C_{75}^5 \cdot C_{15}^1} \approx \frac{1}{21}$$

$$P(5,0) = \frac{C_{14}^1}{C_{75}^5 \cdot C_{15}^1} \approx \frac{1}{18,492,204}$$

$$P(4,0) = \frac{C_5^4 \cdot C_{70}^1 \cdot C_{14}^1}{C_{75}^5 \cdot C_{15}^1} \approx \frac{1}{52,835}$$

$$P(3,0) = \frac{C_5^3 \cdot C_{70}^2 \cdot C_{14}^1}{C_{75}^5 \cdot C_{15}^1} \approx \frac{1}{766}$$

$$P(2,0) = \frac{C_5^2 \cdot C_{70}^3 \cdot C_{14}^1}{C_{75}^5 \cdot C_{15}^1} \approx \frac{1}{34}$$

$$P(1,0) = \frac{C_5^1 \cdot C_{70}^4 \cdot C_{14}^1}{C_{75}^5 \cdot C_{15}^1} \approx \frac{1}{4}$$

$$P(0,0) = \frac{C_5^0 \cdot C_{70}^5 \cdot C_{14}^1}{C_{75}^5 \cdot C_{15}^1} \approx \frac{2}{3}$$

3 Variables and Expectations

3.1 Average and Deviation

Referring back to the founding definition (1), note that the normalization condition $P(A) + P(\bar{A}) = 1$ may be written as a sum over each event index k via

$$1 = \sum_{k=1}^{n} P(A_k) \ , \tag{11}$$

where n is the total number of events or trials. Multiplying by a factor of A_k inside the sum, we arrive at the precise definition of a weighted average of A_k, denoted $\langle A \rangle$ according to

$$\langle A \rangle = \sum_{k=1}^{n} A_k \cdot P(A_k) \ . \tag{12}$$

The above generalizes to calculate the *expectation value* any dependent function $f(A_k)$ as

$$\langle f \rangle = \sum_{k=1}^{n} f(A_k) \cdot P(A_k) \ , \tag{13}$$

where a few common choices for $f(A_k)$ are 1, A_k, or A_k^2.

The *standard deviation* in the quantity $f(A_k)$ is defined as

$$\sigma_f = \sqrt{\sum_{k=1}^{n} \left(f(A_k) - \langle f \rangle\right)^2 \ P(A_k)} \ , \tag{14}$$

which, using the definitions above, readily reduces to

$$\sigma_f = \sqrt{\langle f^2 \rangle - 2 \langle f \rangle \langle f \rangle + \langle f \rangle^2} = \sqrt{\langle f^2 \rangle - \langle f \rangle^2} \, .$$

Problem 24

A six-sided dice that chooses a random number $1 \leq A_k \leq 6$ is tossed in succession to produce $N \gg 1$ events. Calculate the average outcome.

Solution 24

$$\langle A \rangle = \frac{1}{6} + \frac{2}{6} + \frac{3}{6} + \frac{4}{6} + \frac{5}{6} + \frac{6}{6} = \frac{21}{6} = 3.5$$

Problem 25

A six-sided dice that is missing the 2-face but has an extra 4-face is tossed in succession to produce $N \gg 1$ events. Calculate the average outcome.

Solution 25

$$\langle A \rangle = \frac{1}{6} + \frac{0}{6} + \frac{3}{6} + \frac{4 \cdot 2}{6} + \frac{5}{6} + \frac{6}{6} = \frac{23}{6} = 3.83$$

3.2 Continuous Distributions

For a stochastic process that produces outputs A_k in a continuous range instead of a discrete set, the normalization condition

$$\sum_k P(A_k) = 1$$

becomes an infinite sum, limiting to an integral as

$$1 = \sum_{k=1}^{n} P(A_k) = \int_n dP(k) = \int_n \frac{dP(k)}{dk} \, dk = \int_n w(k) \, dk \, . \tag{15}$$

We abbreviate the notation $A_k \to k$ in transition from a discrete sample space to a continuous one. The continuous function $w(k)$ is called the *probability density*, or *probability distribution* (although 'w' stands for *weight*). Specifically, $\omega(k)$ is the probability of an event occurring within a window $[k, k + dk]$.

In light of equation (15), the equations for the weighted average, general expectation value, and standard deviation generalize to:

$$\langle k \rangle = \int_n k \cdot w(k) \, dk \, . \tag{16}$$

$$\langle f \rangle = \int_n f(k) \, w(k) \, dk \, . \tag{17}$$

$$\sigma_f = \sqrt{\int_n (f(k) - \langle f \rangle)^2 \, w(k) \, dk} = \sqrt{\langle f^2 \rangle - \langle f \rangle^2} \tag{18}$$

Just No Ordinary Triangle

Source: Brilliant, Efren Medallo
`https://brilliant.org/`

Problem 26
What is the expected area of a right triangle with a hypotenuse of k whose non-right angles are uniformly distributed over the interval $(0, \pi/2)$?

Solution 26

$$\langle A \rangle = \frac{k^2/2}{\pi/2} \int_0^{\pi/2} \cos\theta \sin\theta \, d\theta = \frac{k^2/2}{\pi/2} \int_0^1 x \, dx = \frac{k^2}{2\pi}$$

Quadrilateral

Source: Canyon Crest Academy, Shay
`http://teachers.sduhsd.net/bshay/Fun%20Problems.pdf`

Problem 27
Divide a given line segment into two other line segments. Then, cut each of these new line segments into two more line segments. What is the probability that the resulting four line segments are the sides of a quadrilateral?

Solution 27
Let the total length be L, and require that no one side be longer than $L/2$. After the initial cut, let the longer segment have length x, and the shorter segment $L - x$. Diving the longer segment at point z (from the start of x), it is required that $z < L/2$ and simultaneously $x - z < L/2$. Therefore, the window of allowed z has width $L/2 - (x - L/2) = L - x$. The normalized probability of an allowed z along x is:

$$P = N \int_{L/2}^{L} \frac{L - x}{x} dx = \frac{(L \ln x - x)\big|_{L/2}^{L}}{L/2} = 2\ln 2 - 1 \approx 38.6\%$$

3.3 Sum of Random Variables

Consider a set $\{A_k\}$ of random (not necessarily independent) variables whose sum is A:

$$A = \sum_{k=1}^{n} A_k$$

In the continuous limit, the normalization condition generalizes to a k-dimensional integral

$$1 = \sum_{k=1}^{n} P(A_k) = \int_n dP(A_k) = \int w(A_1, A_2, \ldots, A_k) \, dA_1 \, dA_2 \ldots dA_k \, . \tag{19}$$

Assuming nothing about the probability distribution $w(A_1, A_2, \ldots, A_n)$, we can easily prove the 'sum of the averages' rule by calculating the expectation value $\langle A \rangle$:

$$\begin{aligned}
\langle A \rangle &= \sum_{k=1}^{n} \int A_k \, w(A_1, A_2, \ldots, A_n) \; dA_1 \, dA_2 \ldots dA_n \\
&= \int_n (A_1 + A_2 + \cdots A_n) \; w(A_1, A_2, \ldots, A_n) \; dA_1 \, dA_2 \ldots dA_n \\
&= \langle A_1 \rangle + \langle A_2 \rangle + \cdots + \langle A_n \rangle \\
\langle A \rangle &= \sum_{k=1}^{n} \langle A_k \rangle
\end{aligned} \tag{20}$$

Strictly translated, the above reads 'the average of the sum is the sum of the averages'.

3.4 Independent Random Variables

Recall the normalization condition for a set of random variables,

$$1 = \int_n w(A_1, A_2, \ldots, A_n) \; dA_1 \, dA_2 \ldots dA_n \, .$$

Generalizing equation (4) that handles discrete independent events, the probability distribution w fragments into individual products:

$$w(A_1, A_2, \ldots, A_n) = w_1(A_1) \cdot w_2(A_2) \cdots w_n(A_n) \tag{21}$$

One immediate result we can write is an analog to equation (20), but for multiplication:

$$\begin{aligned}
\langle A_1 \cdot A_2 \cdots A_n \rangle &= \int (A_1 \cdot A_2 \cdots A_n) \; w(A_1, A_2, \ldots, A_n) \; dA_1 \, dA_2 \ldots dA_n \\
&= \langle A_1 \rangle \cdot \langle A_2 \rangle \cdots \langle A_n \rangle
\end{aligned} \tag{22}$$

3.5 Variance

Consider a set $\{A_k\}$ of n random variables whose sum is A. The square of A is written

$$A^2 = \left(\sum_{i=1}^{n} A_i \right) \left(\sum_{j=1}^{n} A_j \right) = \sum_{k=1}^{n} A_k^2 + \sum_{i \neq j} c_{ij} A_i A_j \, ,$$

where c_{ij} are the binomial coefficients to represent the cross terms. Meanwhile, the square of the average $\langle A \rangle$ comes out to

$$\langle A \rangle^2 = \sum_k \langle A_k \rangle^2 + \sum_{i \neq j} c_{ij} \langle A_i \rangle \langle A_j \rangle$$

In order to calculate $\langle A^2 \rangle$, the independence among A_k must be invoked. Using equation (21) and its consequences, we calculate $\langle A^2 \rangle$ to be

$$\left\langle A^2 \right\rangle = \sum_{k=1}^{n} \left\langle A_k^2 \right\rangle + \sum_{i \neq j} c_{ij} \left\langle A_i \right\rangle \left\langle A_j \right\rangle$$

Taking the difference $\langle A^2 \rangle - \langle A \rangle^2$, the cross terms cancel and we arrive at a simple relation connecting A to its members:

$$\left\langle A^2 \right\rangle - \left\langle A \right\rangle^2 = \sum_{k=1}^{n} \left\langle A_k^2 \right\rangle - \left\langle A_k \right\rangle^2 + \cancel{\sum_{i \neq j} c_{ij} \left\langle A_i \right\rangle \left\langle A_j \right\rangle} - \cancel{\sum_{i \neq j} c_{ij} \left\langle A_i \right\rangle \left\langle A_j \right\rangle}$$

The square root of $\langle A^2 \rangle - \langle A \rangle^2$ is defined as the *variance* in A:

$$Var\left(A\right) = \sqrt{\langle A^2 \rangle - \langle A \rangle^2} = \sqrt{\sum_{k=1}^{n} \langle A_k^2 \rangle - \langle A_k \rangle^2} = \sqrt{\sum_{k=1}^{n} \left(Var\left(A_k\right)\right)^2} \tag{23}$$

3.6 Dispersion

A variation in A, denoted ΔA, is known as *dispersion*, defined as

$$\Delta A = A - \langle A \rangle = \sum_{k=1}^{n} A_k - \langle A_k \rangle \ ,$$

where all A_k are independent. The expectation value $\langle \Delta A \rangle = 0$ clearly, however $\langle \Delta A^2 \rangle$ is interesting:

$$\begin{aligned}
\left\langle \Delta A^2 \right\rangle &= \int_n \left(A_1 + \cdots + A_n - \langle A_1 \rangle - \cdots - \langle A_k \rangle\right)^2 \, w\left(A_1, A_2, \ldots, A_n\right) \, dA_1 \, dA_2 \ldots dA_n \\
&= \sum_{k=1}^{n} \left\langle \left(A_k - \langle A_k \rangle\right)^2 \right\rangle + \cancel{\sum_{i \neq j} \tilde{c}_{ij} \left\langle A_i \right\rangle \left\langle A_j \right\rangle} \\
&= \sum_{k=1}^{n} \left\langle \Delta A_k^2 \right\rangle
\end{aligned}$$

Note that the coefficient $\tilde{c}_{ij}$ resolves to zero for all mixed terms. (If the above relation is difficult to penetrate at a glance, try with only two variables $A_1 = x$ and $A_2 = y$ to derive the $k = 2$ case.)

In the large-n limit, the average $\langle A \rangle$ scales with n, and meanwhile we see $\langle \Delta A^2 \rangle$ also scales with n. The ratio of the RMS dispersion over the average thus tends to zero, as

$$\frac{\sqrt{\langle \Delta A^2 \rangle}}{\langle A \rangle} \approx \frac{1}{\sqrt{n}} \to 0 \, ,$$

telling us that fluctuations in A become negligibly small.

4 Systems and Distributions

4.1 Two-State System

Consider a balanced coin that is tossed to generate n random events resulting in either H(eads) or T(ails). If we are interested in the portion m 'heads' events that occur without the order of events being important, the number of combinations is given by (10), namely

$$C_n^m = \frac{n!}{m!\,(n-m)!} \; .$$

The sum of all C_n^m across the whole range of m, namely from 0 to n, must resolve to the total multiplicity of events, namely 2^n:

$$2^n = \sum_{m=0}^{n} \frac{n!}{m!\,(n-m)!}$$

Evidently, the probability of attaining m events among n trials in any two-state system is

$$P\left(m,n\right) = \frac{1}{2^n}\frac{n!}{m!\,(n-m)!} \; . \tag{24}$$

The combination number C_n^m can be interpreted nicely by spotting the pattern that emerges in trivial cases. A single toss can result in T or H, which we denote

$$\omega_1 = \left(T, H\right) \; .$$

Denoting m as the number of H-events, we write

$$C_1^0 = 1 \qquad\qquad C_1^1 = 1$$

For a game of $n = 2$ tosses, the list of possible events is

$$\omega_2 = \left(TT, TH, HT, HH\right) \; .$$

Again denoting m as the number of H-events, we write

$$C_2^0 = 1 \qquad\qquad C_2^1 = 2 \qquad\qquad C_2^2 = 1$$

Similarly, a game of three tosses has

$$\omega_3 = \left(TTT, TTH, THT, THH, HTT, HTH, HHH\right) \; ,$$

having combinations

$$C_3^0 = 1 \qquad\qquad C_3^1 = 3 \qquad\qquad C_3^2 = 3 \qquad\qquad C_3^3 = 1 \; .$$

The pattern in C_n^m (stand back and look at it) evidently matches the rows of *Pascal's triangle*, which is a map of the coefficients attained when expanding the binomial $(x+y)^n$. The C_n^m are called binomial coefficients for this reason.

Finally, there is yet another intuitive way to generate C_n^m. Begin with n coins, all turned to the T-face. The strategy is to gradually introduce H-events into the system by flipping any of the n coins. Starting slowly, a game of n tosses and zero H-events is unique, thus we have $C_n^0 = 1$. Next, flip any one of the n coins to H, so evidently $C_n^1 = n$. To introduce a second H-event, there are $n-1$ available T-faces to replace. Since the precise order of H-ad T- events does not matter, a factor of $1/2$ must be introduced to avoid double-counting. Raising the H-events to $m = 3$, there are $n-2$ available T-faces to flip, and we require a factor of $1/3$ to avoid triple counting. So far, we may write

$$C_n^3 = n\frac{(n-1)}{2}\frac{(n-2)}{3} \ ,$$

which, by induction for any m matches the anticipated C_n^m formula.

4.2 Binomial Distribution

Consider an *unbalanced* coin having inherent probability p to land on H(eads), and correspondingly $1-p$ to land on T(ails). As a generalized two-state system, a game of n tosses generates the same potential outcomes:

$$\begin{aligned}
\omega_1 &= (T, H) \\
\omega_2 &= (TT, TH, HT, HH) \\
\omega_3 &= (TTT, TTH, THT, THH, HTT, HTH, HHH)
\end{aligned}$$

Of course, the probability $P(m,n,p)$ of generating m Heads-events is no longer proportional to C_n^m alone, as the factors p and $1-p$ must enter the picture. Denoting the modified combination symbol $\tilde{C}$, the two-state analysis generalizes by:

$$\begin{array}{lll}
\tilde{C}_1^0 = 1-p & \tilde{C}_1^1 = p & \\
\tilde{C}_2^0 = (1-p)^2 & \tilde{C}_2^1 = 2\cdot p\,(1-p) & \tilde{C}_2^2 = p^2 \\
\tilde{C}_3^0 = (1-p)^3 & \tilde{C}_3^1 = 3\cdot p\,(1-p)^2 & \\
 & \tilde{C}_3^2 = 3\cdot p^2\,(1-p) & \tilde{C}_3^3 = p^3
\end{array}$$

Evidently, the factors of p and $1-p$ compound into the terms p^m and $(1-p)^{n-m}$, but otherwise this analysis traces that of the two-state system exactly. The multiplicity of the unbalanced coin is therefore

$$\tilde{C}_n^m(p) = \frac{n!}{m!\,(n-m)!}\,(1-p)^{n-m}\,p^m \ .$$

Dividing by the total multiplicity 2^n, we find the probability of attaining m events among n trials in a p-biased two-state system to be

$$P(m,n,p) = \frac{1}{2^n}\frac{n!}{m!\,(n-m)!}\,(1-p)^{n-m}\,p^m \ ,$$

also known as the *binomial distribution*.

To proceed, define a random variable z_k that is equal to one if the event H with weight p occurs in the k-th trial, and is equal to zero otherwise. The average value of z_k is then

$$\langle z_k \rangle = P(H)\, z(H) + P(T)\, z(T) = p \cdot 1 + (1-p) \cdot 0 = p \,,$$

and, simply enough, the average of z_k^2 reads

$$\left\langle z_k^2 \right\rangle = p \cdot 1^2 + (1-p) \cdot 0^2 = p \,.$$

The standard deviation in z, denoted σ_z, is evidently

$$\sigma_z = \sqrt{\left\langle z_k^2 \right\rangle - \langle z_k \rangle^2} = \sqrt{p - p^2} = \sqrt{p(1-p)} \,.$$

Next, note that the number m of H-events among the n independent trials is the sum

$$m = \sum_{k=1}^{n} z_k \,,$$

implying

$$\left\langle \Delta m^2 \right\rangle = \sum_{k=1}^{n} \left\langle \Delta z_k^2 \right\rangle = n \left\langle \Delta z^2 \right\rangle \,,$$

or, in tighter notation for large-n systems,

$$\sigma_m = \sqrt{n\sigma_z^2} = \sqrt{np(1-p)} \,.$$

Goals in Netball

Source: MyMathforum
`http://mymathforum.com/`

Problem 28

Monique is practicing netball. She knows from past experience that the probability of her making any one shot is 70%. Her coach has asked her to keep practicing until she scores 50 goals. How many shots would she need to attempt to ensure that the probability of making at least 50 shots is more than 99%?

Solution 28

This problem is analogous to flipping a weighted coin with bias p. The multiplicity of scoring k shots in N tosses is

$$\Omega(k, N, p) = \frac{N!}{k!\,(N-k)!} (1-p)^{N-k} p^k \,,$$

where summing over k gives the cumulative distribution, from which we pick out the correct N by computer:

$$99\% = \sum_{k=50}^{N} \frac{N!}{k!\,(N-k)!} .3^{N-k} .7^k \qquad \rightarrow \qquad N = 86$$

4.3 Gaussian Distribution

Recall that the probability of generating k results among n total trials in a two-state system is given by

$$P(k,n) = \frac{1}{2^n}\frac{n!}{k!\,(n-k)!} = \frac{1}{2^n}\frac{n!}{\left(\frac{n}{2}+k\right)!\left(\frac{n}{2}-k\right)!}$$

In the large-k limit, where k becomes a continuous variable, it makes sense to describe the system solely in terms of expectation values and their deviations, a notion formally called the *central limiting theorem*. Here we develop this idea on a two-state system to derive a central equation in probability theory called the *Gaussian distribution*.

To proceed in the large n-limit, we deploy Stirling's approximation for large numbers

$$\ln(n!) \approx n\ln(n) - n + \ln\left(\sqrt{2\pi n}\right) \qquad\qquad n! \approx \left(\frac{n}{e}\right)^n\sqrt{2\pi n}\,,$$

and the probability density reduces to

$$w(k) \approx e^{-2k^2/n}\sqrt{\frac{2}{\pi n}}\,.$$

The result $w(k)$ is the famed normalized Gaussian distribution centered at $k=0$. Introducing a nonzero shift of base-point value a, the generalized equation is

$$w(k) = e^{-2(k-a)^2/n}\sqrt{\frac{2}{\pi n}}\,. \tag{25}$$

Using Gaussian integrals, the average values and standard deviation can be calculated:

$$\begin{aligned}\langle k\rangle &= \int_n k\cdot w(k)\;dk = a\\ \left\langle k^2\right\rangle &= \int_n k^2\cdot w(k)\;dk = \frac{n}{4}+a^2\\ \sigma_k &= \sqrt{\langle k^2\rangle - \langle k\rangle^2} = \sqrt{\frac{n}{4}}\end{aligned}$$

4.4 Poisson Distribution

Imagine trying to count the number of water molecules that pass a point in a river flowing at average speed v. Over time interval t, the average molecule count is directly proportional to vt. To reduce notational clutter, let us ignore the proportionality constant and take vt as a dimensionless quantity. Due to local random fluctuations in the river, an actual measurement would never precisely land on vt, but instead on an interval surrounding vt. Naturally we wonder, what is the time-varying probability $P_k(t)$ that k molecules are measured over the interval t?

To begin, partition the elapsed time t into n identical bins of width Δt such that $\Delta t \to 0$, and observe that each $P_k(\Delta t)$ relates to its $k-1$ and $k+1$ neighbors as

$$\lim_{\Delta t\to 0} P_0(\Delta t) \gg P_1(\Delta t) \gg P_2(\Delta t) \gg P_3(\Delta t) \gg \cdots\,,$$

which means it's more likely to measure few molecules in a small Δt-interval as opposed to many. We may proceed using weighted two-state analysis, wherein a Δt-interval may either be unfilled with zero molecules, or filled with one or more molecules. Borrowing the apparatus developed previously, we write

$$P\left(k,n,v\Delta t\right)=\frac{n!}{k!\left(n-k\right)!}\left(1-v\Delta t\right)^{n-k}\left(v\Delta t\right)^{k}\ ,$$

where n and k are integers. Substituting $t=n\Delta t$, we have

$$P\left(k,n,vt\right)=\frac{\left(vt\right)^{k}}{k!}\left(\frac{n!}{\left(n-k\right)!}\frac{1}{n^{k}}\right)\left(1-\frac{vt}{n}\right)^{n-k}\ .$$

In the large-n limit, the approximations

$$\frac{n!}{\left(n-k\right)!}\approx n^{k}\qquad\qquad\left(1-\frac{vt}{n}\right)^{n-k}\approx e^{-vt}$$

are valid, and re-casting vt as a dimensionless variable q lands us at the anticipated *Poisson distribution*:

$$P_{k}\left(q\right)=\frac{q^{k}}{k!}e^{-q}\tag{26}$$

Summing over the variable k tells us $P_k\left(t\right)$ is already normalized:

$$\sum_{k=0}^{\infty}\frac{q^{k}}{k!}e^{-q}=e^{-q}\left(\sum_{k=0}^{\infty}\frac{q^{k}}{k!}\right)=e^{-q}e^{q}=1$$

With $P_k\left(t\right)$ on hand, we may calculate $\langle k\rangle$, $\langle k^2\rangle$, and the standard deviation:

$$\begin{aligned}\langle k\rangle&=\sum_{k=0}^{\infty}k\frac{q^{k}}{k!}e^{-q}=e^{-q}\sum_{k=1}^{\infty}\frac{q^{k}}{\left(k-1\right)!}=e^{-q}\sum_{p=0}^{\infty}\frac{q^{(p+1)}}{p!}=e^{-q}qe^{q}=q\\\left\langle k^{2}\right\rangle&=\sum_{k=0}^{\infty}k^{2}\frac{q^{k}}{k!}e^{-q}=e^{-q}q\sum_{p=0}^{\infty}\left(p+1\right)\frac{q^{p}}{p!}=q+q^{2}\\\sigma_{k}&=\sqrt{q^{2}+q-q^{2}}=\sqrt{q}\end{aligned}$$

5 Appendix

5.1 Stirling's Approximation

Factorials of large numbers obey an approximation that eliminates the factorial (!) symbol called *Stirling's approximation*:

$$\ln\left(n!\right)\approx n\ln\left(n\right)-n+\ln\left(\sqrt{2\pi n}\right)\qquad\qquad n!\approx\left(\frac{n}{e}\right)^{n}\sqrt{2\pi n}\tag{27}$$

To derive Stirling's approximation, begin with the gamma function identity

$$\Gamma(z) = (z-1)! = \int_0^\infty t^{z-1} e^{-t}\, dt$$

to write

$$n! = \int_0^\infty e^{-x} x^n \, dx \, .$$

Using plots of the sharply-peaked function $e^{-x}x^n$, or by calculating its derivative, one may easily argue that the approximation $x \approx n + \epsilon$ holds for $n \gg \epsilon$. Following this, we find

$$\begin{aligned} e^{-x}x^n &\approx e^{-(n+\epsilon)} (n+\epsilon)^n \\ \ln\left(e^{-x}x^n\right) &\approx -n - \epsilon + n \ln(n+\epsilon) \ , \end{aligned}$$

where the expansion for natural logarithms

$$\ln(1+y) = y - \frac{y^2}{2} + \frac{y^3}{3} - ... \qquad\qquad y \ll 1$$

leads us to

$$n! \approx \int_0^\infty \exp\left(n \ln n - n - \frac{\epsilon^2}{2n} + \cancel{\frac{\epsilon^3}{3n^2}} - ...\right) d\epsilon \, ,$$

noting that orders of ϵ greater than two are neglected. The above resolves to

$$n! \approx e^{n \ln n - n} \int_0^\infty e^{-\epsilon^2/2n}\, d\epsilon \, ,$$

and since the integrand is essentially zero away from the peak, we extend the integration limit from 0 to $-\infty$. This transforms the final calculation to a Gaussian integral, obeying

$$\int_{-\infty}^\infty e^{-ax^2}\, dx = \sqrt{\frac{\pi}{a}} \, ,$$

which reproduces equation (27):

$$n! \approx \left(\frac{n}{e}\right)^n \sqrt{2\pi n} \, .$$

5.2 Copernicus Method

Source: Futility Closet
`https://futilitycloset.com/`

Princeton astrophysicist J. Richard Gott was visiting the Berlin Wall in 1969 when a curious thought occurred to him. His visit occurred at a random moment t in the wall's total lifespan T, so it seemed reasonable to assume that there was a 50% percent chance that he was observing it in the middle two quarters of its lifetime. At the time, the wall was 8 years old, so Gott concluded that there was a 50% chance that it would last for at least

$8 \times 1/3$ additional years but fewer than 8×3 additional years. The wall came down in 20 years later in 1989.

The 'visit time' t years is, in the two extremes, either (a) one quarter of the wall's lifetime, or (b) three quarters of the wall's lifetime. For case (a), the remaining lifetime is 3 more counts of t years, giving $T_{max} = t + 3t = 4t$ years. For case (b), the t years that have elapsed correspond to $t = (3/4) \times T_{min}$, or $T_{min} = t \times (4/3)$. Inserting $t = 8$, we find

$$T_{max} = 4 \times 8\,yr = 32\,yr \qquad\qquad T_{min} = \frac{4}{3} \times 8\,yr \approx 10.67\,yr$$

or

$$T_{max} \approx 1993 \qquad\qquad T_{min} \approx 1971\,.$$

Generalization

Next we generalize the above result for N bins instead of four. The time t is the ratio of either (a) $1/N$ of the wall's lifetime, or (b) $(N-1)/N$ of the wall's lifetime. For case (a), the remaining lifetime is $(N-1)$ more counts of t years, giving $T_{max} = Nt$ years. For case (b), the t years that have elapsed correspond to $t = ((N-1)/N) \times T_{min}$, or $T_{min} = t \times (N/(N-1))$. In summary, we have

$$T_{max} = Nt \qquad\qquad T_{min} = \frac{Nt}{N-1}\,.$$

Of course, the window defined by $T_{max} - T_{min}$ no longer corresponds to a probability of 50%, but must be adjusted to

$$p\,(N) = \frac{N-2}{N}\,.$$

Example

Suppose you encounter a man for the first time in the 42^{nd} year of his life. Determine the upper and lower bounds of the interval in which he has a 33% chance of expiring (in total years). Repeat for 66% and 75%.

$$T_{max(33\%)} = 3 \times 42 = 126 \qquad\qquad T_{min(33\%)} = \frac{3}{2} \times 42 = 63$$

$$T_{max(66\%)} = 6 \times 42 = 252 \qquad\qquad T_{min(66\%)} = \frac{6}{5} \times 42 = 50.4$$

$$T_{max(75\%)} = 8 \times 42 = 336 \qquad\qquad T_{min(75\%)} = \frac{8}{7} \times 42 = 48$$