Bernoulli's Theorem

Let R be the number of ways of getting a success in a trial and let S be the number of ways of getting a failure in a trial. Let T = R + S. Let N and C be positive integers Then in NT trials for a given C if N is sufficiently large it will be at least C times as likely to get a number of successes in the range NR-N through NR+N than outside that range.

Comment

If the number of successes in NT trials is divided by NT we get the relative frequency of successes in NT trials.

So if the number of successes in NT trials is in the range, the relative frequency will be in the range R/T -1/T through R/T + 1/T where R/T is the probability of success in a single trial. For example, if R is 30 and S is 20 we get a relative frequency range of 3/5 – 1/50 through 3/5 + 1/50 .

Bernoulli derived a formula to determine the value needed for NT which would guarantee that it would be at least C times as likely for the number of successes to lie in the range NR-N through NR+N than outside that range.

With R = 30, S = 20, and C = 1000, Bernoulli's formula gives an NT of 25,550.

In this paper, we prove that an NT of 15,200 is sufficient for Bernoulli's example. This is a 40.5% reduction in NT and the method used is much simpler than Bernoulli's method. Bernoulli's method is shown in my paper <u>Bernoulli's Scholium.</u> My method begins with my lemma 7.

Proof of theorem:

PART ONE

Let W_i be the number of ways of getting exactly i successes in NT trials.

Lemma 1 $W_{K+1}/W_K = \frac{(NT-K)R}{(K+1)S}$

From Lemma 1 we see that as the number of successes K increases, the ratio W_{K+1}/W_K decreases.

Proof of lemma:

P(K), the probability of getting exactly K successes in NT trials is:

$$\frac{NT(NT-1)\ldots(NT-K+1)}{K(K-1)\ldots1}\left(\frac{R}{T}\right)^{K}\left(\frac{S}{T}\right)^{NT-K}$$

If K is replaced by K+1 in this formula, there will be an extra factor of NT-K in the numerator, an extra factor of K+1 in the denominator, an extra factor of R in the numerator and one less factor of S in the numerator. .

So $P(K+1) = \frac{(NT-K)R}{(K+1)S}P(K)$ and $\frac{P(K+1)}{P(K)} = \frac{(NT-K)R}{(K+1)S}$

Since there are T things that can happen in a given trial, then in NT trials the total number of ways things can happen is T^{NT}.

So since $\frac{P(K+1)}{P(K)} = \frac{\frac{W_{K+1}}{T^{NT}}}{\frac{W_K}{T^{NT}}} = \frac{W_{K+1}}{W_K}$ lemma 1 is proved.

Lemma 2: W_{NR} is greater than all the other W's and the W's get smaller as their subscripts decrease from NR and as they increase from NR.

Proof

From lemma 1, $\frac{W_{NR+1}}{W_{NR}} = \frac{(NT-NR)R}{(NR+1)S} = \frac{NRS}{NRS+S} < 1$

So $W_{NR+1} < W_{NR}$. $\frac{W_{NR}}{W_{NR-1}} = \frac{(NT-NR+1)R}{(NR-1+1)S} = \frac{NRS+R}{NRS} > 1$

So $W_{NR} > W_{NR-1}$. So since the ratios are decreasing as the subscripts increase, this means the ratios are all greater than 1 from $\frac{W_1}{W_0}$ through $\frac{W_{NR}}{W_{NR-1}}$. So $W_0<W_1<\ldots<W_{NR}$

And since $\frac{W_{NR+1}}{W_{NR}} < 1$, all the ratios from $\frac{W_{NR+1}}{W_{NR}}$ through $\frac{W_{NT}}{W_{NT-1}}$ will be less than 1. So $W_{NR} > W_{NR+1} > \ldots > W_{NT}$.

Lemma 3: The ratio

$(W_{NR+1} + W_{NR+2} + \ldots + W_{NR+N}) / (W_{NR+N+1} + \ldots + W_{NT})$ (1)

is greater than or equal to:

$(W_{NR+1} + W_{NR+2} + \ldots + W_{NR+N}) / (S-1)(W_{NR+N+1} + \ldots + W_{NR+2N})$. (2)

This is true because from lemma 2 we know that the terms in the denominator of (1) are decreasing, so the sum of the first N terms is greater than the sum of any subsequent N terms and since there are (S–1)N terms in the denominator of (1), denominator of (2) $\geq$ denominator of (1). So the ratio (2) $\leq$ ratio (1) (Equality occurs when S = 2 because then NT = NR + 2N and so there are no subsequent N terms.)

Lemma 4:

$$\frac{W_{NR}}{W_{NR+N}} < \frac{W_{NR+1}}{W_{NR+N+1}} < \frac{W_{NR+2}}{W_{NR+N+2}} < \ldots < \frac{W_{NR+N}}{W_{NR+2N}}$$

Proof: Let A be a nonnegative integer, then using lemma 1 we get

$$\frac{W_{NR+N+A+1}}{W_{NR+N+A}} < \frac{W_{NR+A+1}}{W_{NR+A}}$$

because NR + N + A > NR + A.

Multiplying both sides of the inequality by $\frac{W_{NR+A}}{W_{NR+N+A+1}}$ we get:

$$\frac{W_{NR+A}}{W_{NR+N+A}} < \frac{W_{NR+A+1}}{W_{NR+N+A+1}}$$

Substituting 0,1, 2 ,3, …,N-1 for A we get the inequalities of the lemma.

Lemma 5

Let $A_1, A_2, \ldots, A_N$ and $B_1, B_2, \ldots, B_N$ be positive integers then if $A_1/B_1 < A_2/B_2 < A_3/B_3 < \ldots < A_{N-1}/B_{N-1} < A_N/B_N$ then $A_1/B_1 < (A_2 + A_3 + \ldots + A_N)/(B_2 + B_3 + \ldots + B_N)$.

proof of Lemma 5

From the chain of inequalities, it follows that $\frac{A_1}{B_1} < \frac{A_i}{B_i}$ where i > 1

So $A_1B_i < B_1A_i$ for $i > 1$

This gives the following inequalities:

$A_1B_2 < B_1A_2$

$A_1B_3 < B_1A_3$

* *

* *

$A_1B_N < B_1A_N$

Adding the terms on the left side of the inequalities and adding the terms on the right side of the inequalities gives the following inequality:

$A_1(B_2 + B_3 + B_4 + \ldots + B_N) < B_1(A_2 + A_3 + A_4 + \ldots + A_N)$

or $A_1/B_1 < (A_2 + A_3 + \ldots + A_N)/ (B_2 + B_3 + \ldots + B_N)$

Lemma 6 W_{NR}/W_{NR+N} is less than:

$(W_{NR+1} + W_{NR+2} + \ldots + W_{NR+N})/ (W_{NR+N+1} + \ldots + W_{NR+2N})$

This follows from lemmas 4 and 5.

Lemma 7

The ratio W_{NR}/W_{NR+N} can be made as large as desired by making N sufficiently large.

Proof

By substituting NR and NR+N for K in the formula for the probability of getting exactly K successes in NT trials , we get:

$$P(NR) = \frac{NT(NT-1)***(NT-NR+1)}{NR(NR-1)***1}\left(\frac{R}{T}\right)^{NR}\left(\frac{S}{T}\right)^{NT-NR}$$

$$P(NR+N) = \frac{NT(NT-1)***(NT-NR-N+1)}{(NR+N)(NR+N-1)***1}\left(\frac{R}{T}\right)^{NR+N}\left(\frac{S}{T}\right)^{NT-NR-N}$$

So, since $W_{NR}/W_{NR+N} = (W_{NR}/T^{NT})/(W_{NR+N}/T^{NT}) = P(NR)/P(NR+N)$ we get:

$$W_{NR}/W_{NR+N} = \frac{NR+N}{NS} * \frac{NR+N-1}{NS-1} * \ldots * \frac{NR+1}{NS-N+1} * \frac{S^N}{R^N}$$

$$= \frac{NR+1}{NS-N+1} * \frac{NR+2}{NS-N+2} * \ldots * \frac{NR+N}{NS} * \frac{S^N}{R^N}$$

$$= \frac{NRS+S}{NRS-NR+R} * \frac{NRS+2S}{NRS-NR+2R} * \ldots * \frac{NRS+NS}{NRS}$$

This is a product of N fractions

.

Notice that each numerator is obtained from the previous numerator by adding S to it and each denominator is obtained from the previous denominator by adding R to it.

The first fraction $\frac{NRS + S}{NRS - NR + R}$ is itself obtained from

$\frac{NRS}{NRS - NR}$ by adding S to the numerator and R to the denominator. Moving from left to right, each fraction will be closer to S/R than the previous one, so if NRS/(NRS-NR) < S/R, the fractions will be increasing and NRS/(NRS-NR) will be smaller than the other fractions. If NRS/(NRS-NR) >S/R the fractions will be decreasing and (NRS+NS)/NRS will be the smallest fraction.

If NRS/(NRS-NR) = S/R, all the fractions will be the same.

(NRS+NS)/NRS = (R+1)/R and NRS/(NRS-NR) = S/(S-1)

Since we have a product of N fractions, W_{NR}/W_{NR+N} will be greater than or equal to the smaller of $((R+1)/R)^N$ or $(S/(S-1))^N$

So since (R+1)/R and S/(S-1) are both constants greater than 1, then by making N sufficiently large, W_{NR}/W_{NR+N}

will be as large as desired.

Lemma 8

The ratio $(W_{NR+1} + W_{NR+2} + \ldots + W_{NR+N})/(W_{NR+N+1} + \ldots + W_{NT})$ will be larger than C if N is sufficiently large.

This follows from Lemma 7, Lemma 6, and Lemma 3.

By lemma 6, we know that W_{NR}/W_{NR+N} is less than

$$(W_{NR+1} + W_{NR+2} + \ldots + W_{NR+N})/ (W_{NR+N+1} + \ldots + W_{NR+2N})$$

Combining this with lemma 3 we see that if

$W_{NR}/W_{NR+N} >= C(S-1)$ lemma 8 will be true and lemma 7 guarantees we can choose an N so that

$W_{NR}/W_{NR+N} >= C(S-1)$.

As an example suppose R = 30 and S = 20 and C = 1000,

then C(S-1) = 19,000 , T =50,

(R+1)/R = 31/30 and S/(S-1) = 20/19. Since 31/30 is smaller than 20/19 we use logarithms to find the value of N such that $(31/30)^N =$ 19,000. Log(19,000)/Log(31/30) = 300.465 which rounded up to an integer is 301. NT is 301x50 = 15,050.

So if NT is greater than or equal to 15,050,

$(W_{NR+1} + W_{NR+2} + \ldots + W_{NR+N})/(W_{NR+N+1} + \ldots + W_{NT}) > 1,000$

PART TWO

To complete the proof we need to show that the ratio

$(W_{NR-1} + W_{NR-2} + \ldots + W_{NR-N})/(W_{NR-N-1} + W_{NR-N-2} + \ldots + W_0)$

will be larger than C if N is sufficiently large.

We proceed the same way as before.

- There are (R-1)N terms in the above denominator, so

  $(W_{NR-1} + W_{NR-2} + \ldots + W_{NR-N})/ (W_{NR-N-1} + W_{NR-N-2} + \ldots + W_0)$

  is greater than or equal to:

  $(W_{NR-1} + W_{NR-2} + \ldots + W_{NR-N})/(R-1)(W_{NR-N-1} +\ldots+ W_{NR-2N})$.

- $W_{NR}/W_{NR-N} < W_{NR-1}/W_{NR-N-1} < W_{NR-2}/W_{NR-N-2} < \ldots < W_{NR-N}/W_{NR-2N}$

- W_{NR}/W_{NR-N} is less than

  $(W_{NR-1} + W_{NR-2} +\ldots+ W_{NR-N})/(W_{NR-N-1} +W_{NR-N-2} +\ldots+W_{NR-2N})$

  - W_{NR}/W_{NR-N} is greater than or equal to the smaller of

    $(R/(R-1))^N$ or $((S+1)/S)^N$, so by making N sufficiently large ,

  W_{NR}/W_{NR-N} will be greater than C(R-1) and then

$(W_{NR-1} + W_{NR-2} + \ldots + W_{NR-N})/(W_{NR-N-1} + W_{NR-N-2} + \ldots + W_0)>C$

In the example C(R-1) is 29,000, T is 50, R/(R-1) is 30/29 and (S+1)/S is 21/20. 30/29 < 21/20, so by taking logarithms we find that when N is 304 , $(30/29)^N$ will be larger than 29,000.
So if NT is greater than or equal to 304 x 50 = 15,200

$(W_{NR-1} + W_{NR-2} + \ldots + W_{NR-N})/(W_{NR-N-1} + W_{NR-N-2}+\ldots+ W_0)> 1,000$

In part one, we had 15,050 trials, so if we use 15,200

both

$(W_{NR-1} + W_{NR-2} + \ldots + W_{NR-N})/(W_{NR-N-1} + W_{NR-N-2}+\ldots+ W_0)$

and

$(W_{NR+1} + W_{NR+2} + \ldots+W_{NR+N})/(W_{NR+N+1}+ \ldots+W_{NT})$

will be greater than 1,000.

Lemma 9

Let A, B, C, D, E be positive integers. If

$(A/B) > C$ and $(D/E) > C$ then $(A+D)/(B+E) > C$.

Proof

$A>BC$ and $D>CE$, so $A+D > BC+CE$. So $(A+D)/(B+E)>C$.

Applying this lemma to the two ratios before the lemma, we find that when NT = 15,200, the number of ways of getting a number of successes in the range NR-N through NR+N(excluding the number of ways of getting exactly NR successes) will be greater than1000 times the number of ways of getting a number of successes outside the range. Adding the number of ways of getting exactly NR successes to the numerator only increases the ratio which was already greater than 1000.

End of proof.

Comment

In order to make 1/T as small as desired, if r and s are the actual number of ways of getting a success or failure on a single trial, we can make, R=mr and S=ms , where m is a positive integer chosen to make 1/T as small as desired. For example, Bernoulli said that if r=3 and s=2, we could make R=30 and S=20 . This would give us a 1/T of 1/50 instead of 1/5.

The value of NT needed to guarantee that it would be at least C times more likely that the relative frequency of successes lies in the range 3/5-1/50 through 3/5+1/50 would be the same whether there was actually 3 ways of getting a success and 2 ways of getting a failure or whether there was 30 ways of getting a success and 20 ways of getting a failure because in either case the probability of getting a success is 3/5, but in order to calculate the NT needed for a range of +- 1/50, R=30 and S=20 have to be used.

Final comment

This paper is a modernization of Bernoulli's proof but Lemma 7 which says that $\frac{W_{NR}}{W_{NR+N}}$ can be made as large as desired by making N sufficiently large is completely different . In Jacob Bernoulli's Ars Conjectandi, he does this in a section called Scholium (explanatory comment). In calculating $\frac{W_{NR}}{W_{NR+N}}$, which he does before he gets to the scholium, the factors in the numerator of the fraction he arrives at are arranged in the reverse order than what I arrived at. So he had to use a different method than mine.

My method is simpler and for Bernoulli's example, it gives an NT of 15,200 instead of 25,550.

This paper is the result of reading Maistrov's outline of Bernoulli's proof in his book, Probability Theory A Historical Sketch

Bernoulli uses the notation M/L instead of $\frac{W_{NR}}{W_{NR+N}}$ and after

presenting his fraction which is arranged differently than the one in my Lemma 7, he gives an argument involving infinity to show that M/L can be made as large as desired by making N sufficiently large. Maistrov's outline mentions that for those who were not satisfied with this argument, Bernoulli gave another argument. This was the one in the Scholium. Maistrov's outline doesn't give the other argument and since I wasn't satisfied with the infinity argument, I decided to see if I could come up with my own argument. The result is my Lemma 7. Maistrov's presentation doesn't say anything about Bernoulli's example or about a formula.

So it was a nice bonus to find out later that using my method, NT gets reduced by 40.5% . All I was trying to do was get a real proof for Bernoulli's lemma 4 (my lemma 7). Bernoulli's infinity argument was unacceptable to me. Reading Maistrov, I didn't even know there was a formula. This wasn't about improving formulas, It was about getting a real proof. Maistrov didn't show the Scholium, if he did, I would have been satisfied because that's where the real proof is.

Daniel Daniels updated 12 /2/2022