The Shooting-Room Paradox and Conditionalizing on Measurably Challenged Sets
Author(s): Paul Bartha and Christopher Hitchcock
Source: *Synthese*, Vol. 118, No. 3 (1999), pp. 403-437
Published by: Springer
Stable URL: http://www.jstor.org/stable/20118152

PAUL BARTHA and CHRISTOPHER HITCHCOCK

THE SHOOTING-ROOM PARADOX AND CONDITIONALIZING ON MEASURABLY CHALLENGED SETS

ABSTRACT. We provide a solution to the well-known "Shooting-Room" paradox, developed by John Leslie in connection with his Doomsday Argument. In the "Shooting-Room" paradox, the death of an individual is contingent upon an event that has a 1/36 chance of occurring, yet the relative frequency of death in the relevant population is 0.9. There are two intuitively plausible arguments, one concluding that the appropriate subjective probability of death is 1/36, the other that this probability is 0.9. How are these two values to be reconciled? We show that only the first argument is valid for a standard, countably additive probability distribution. However, both lines of reasoning are legitimate if probabilities are non-standard. The subjective probability of death rises from 1/36 to 0.9 by conditionalizing on an event that is not measurable, or whose probability is zero. Thus we can sometimes meaningfully ascribe conditional probabilities even when the event conditionalized upon is not of positive finite (or even infinitesimal) measure.

In a dark corner of a hotel bar, several figures sporting name tags are hunched around a small table covered in empty beer bottles. One of the figures is writing furiously on a napkin. The conversation is excited and confused, as many try to speak at once. The individual words that do escape from the white noise of the bar are ominous: "... executioner ...", "... doomsday ...", "... double sixes and you die!"

Any participant in a recent philosophy conference is likely to have encountered a scene much like this. Radical philosophers plotting the overthrow of civilization? No. They are merely curious thinkers testing their mettle on a paradox of probability: the shooting-room paradox. In Section 1.1 of this paper, we will describe this paradox and present the well-known solution to one part of it. In Section 1.2, we will sketch the solution to the more intransigent part of the paradox. The subsequent sections provide the mathematical details upon which the solution depends. We believe that these details are of interest in their own right, and may have applications in a number of philosophical problems.

Synthese **118:** 403–437, 1999.

1. INTRODUCTION

1.1. *The Ballad of George and Tracy*

Imagine that our protagonist, George, has been called to a place known throughout the world as the Shooting-Room. There are dire consequences for those who do not heed the call. As he enters the room, he reads a portentous sign above the door:

> Abandon all hope, you who enter this room!
> Well, not quite all hope – and here's why:
> You've a 1/36 chance of meeting your doom,
> Yet 0.9 of those entering will die!

After George enters, a sinister figure at the front of the room – the "Executioner" – announces that he will roll two ordinary six-sided dice. If the result is double sixes, George (and anyone else in the room with George) will be executed. This explanation seems to entail that George has a one in thirty-six chance of dying. Puzzle: how can the executioner's explanation be reconciled with the statement that 90% of those entering will die?

Here are some hints: (1) The world in which George lives has a countably infinite population. (2) The shooting room has movable walls, so that it can be expanded to arbitrarily large size. (3). The executioner is immortal. (4) The last sentence of the verse above the door is not guaranteed to be true: there is a possibility that it will not be true, but the probability of this eventuality coming to pass is zero (just as it is possible that a coin tossed infinitely many times will never land heads, although there is zero probability that this will happen).

Still stumped? Here is the solution: the shooting-room game is played in rounds. In each round, a larger number of people is called to the shooting room. In the first round, only one person is called; in the second round nine; in the third round ninety; and in each successive round, ten times the number of people called for the preceding round. The game ends only if and when double six is rolled in a given round, in which case all those in the chamber in the last round are executed.[1] There is a probability of one that double sixes will eventually be rolled, and when that occurs, exactly ninety percent of those who have ever entered the room will be present in the final, fatal, round of the game.[2]

This puzzle contains an important moral about the relationship between single-case probabilities and frequencies: we can only expect the latter to conform to the former when (1) individual events (in this case the execution or acquittal of those who enter the shooting-room) are probabilistically independent; and (2) stopping times are probabilistically independent of frequencies. The latter restriction presents genuine problems in drug testing, for example, when ethical considerations mandate the continuation of treatments that appear beneficial, and the termination of treatments that do not.

The apparent paradox was resolved by noting that single case probabilities and predicted frequencies need not coincide. We may recast the paradox, however, in such a way that the values 1/36 and 0.9 both appear to be rational subjective probabilities. Here's how. Suppose that George enters the shooting-room as before, and that he understands the mechanism of the game completely. What should be George's subjective probability for the proposition that he is executed, given that he has been selected to enter the chamber? The answer must be 1/36, for George knows that he will die if and only if the result of a throw of two fair dice is double sixes. Assuming that he can compute that the chance of this outcome is 1/36, that he has no other relevant information, and that his subjective probabilities are guided by his beliefs about chances,[3] his subjective probability should be 1/36.

Now suppose that George's mother, Tracy, also understands the mechanism of the game, knows that George was selected to enter the chamber, and hears the news that the game has ended. She does not know, however, whether George survived. What should be her subjective probability for George to have died? This time, a reasonable answer is 0.9. Tracy knows that a finite number of people entered the room, George among them. Of those people, 90% died (or perhaps 100%, if the game ended on the first round). Tracy has no additional information that is relevant to whether or not George was one of those who participated in the final round, and hence died, so Tracy's subjective probability that George dies should be 0.9.

George and Tracy have exactly the same information, namely that George participated in the game. Neither knows on which round George was selected (we may assume that George was blindfolded, although in fact this does not matter); neither knows how the dice came up on George's round. How then can the probabilities for George and Tracy be so different?

1.2. *Sketch of the Solution*

The previous paragraph contains a little white lie. George and Tracy do not have exactly the same information: Tracy knows that the game has ended, but George does not. On the one hand, this seems crucial to the formulation of the paradox: when George phones Tracy to tell her that he is in the room, and that the executioner is about to roll the dice, it seems that her probability for George's demise (before she learns that the game has ended) should be only 1/36, the same as George's. On the other hand, it is hard to see how the additional information that the game has come to an end could matter: given that both George and Tracy understand the mechanism of the game, they should both assign probability one to the proposition that it eventually ends. Conditioning on a proposition with probability one does not change one's probability assignments, so this extra piece of information could hardly have raised Tracy's subjective probability from 1/36 to 0.9.

It will be helpful, at this point, to make some additional assumptions about the manner in which individuals are selected to participate in the shooting-room game. Assume that each member of the population is assigned a distinct natural number at random: we will call these numbers 'draft' numbers.[4] These numbers are known only to the executioner; however, people know that they will be called no more than once. Individuals are called to the shooting room in the order of their draft numbers. Thus the first group to enter the room will consist of individual number 1; the second group will consist of individuals 2–10; then 11–100 and so on. In general, $9 \cdot 10^{n-2}$ people are selected at round n for $n > 1$. Before George is called to participate in the game, both George and Tracy have prior probabilities for propositions of the form: George has been assigned draft number i. Let us assume that George and Tracy have the same priors: a difference in their posterior probabilities (for George to die, upon learning that he participates in the game) would hardly constitute a paradox if they began with different priors.

Suppose that the subjective probability function initially shared by George and Tracy is standard and countably additive. It follows that they cannot assign uniform probabilities to propositions of the form: George has draft number i. The sum of these prior probabilities must be equal to one, since the probability that *George has some draft number* is one. For example, it could be that the probability that George has draft number i is 2^{-i}, since $1/2 + 1/4 + 1/8 + \cdots$ sums to 1. However, the probability that George has draft number i could not be equal to any constant k. For if $k = 0$, then the 'probability' that *George has some draft number* is $0 + 0 + 0 + \cdots = 0$; whereas if $k > 0$, the 'probability' that *George has some draft number* is $k + k + k + \cdots = \infty$. So the prior probabilities that George has draft number i must converge to 0 as i approaches ∞.

In the case of a countably additive probability function, then, the argument for assigning a value of 0.9 to Tracy's probability that George dies is fallacious. In particular: the statement, "Tracy has no additional information that is relevant to whether or not George was one of those who participated in the final round, and hence died", is false. Tracy's priors provide such information: of necessity, they will be biased toward early draft picks. In Section 2, we demonstrate that no matter what Tracy's initial prior probability is, so long as it is countably additive, it must be biased toward early draft numbers in exactly the manner necessary for her posterior probability for George's death to equal 1/36.

De Finetti took this sort of case to provide an argument against the requirement of countable additivity for probabilities.[5] He reasoned that it ought to be possible to have a lottery in which a countably infinite number of tickets is issued, and that a rational agent should not be prohibited from assigning equal probability to each ticket's winning. Presumably, the original intuition supporting Tracy's assignment of 0.9 was based on the assumption that her beliefs about the draft lottery were of this sort. In Sections 3 and 4, we use nonstandard analysis to construct a probability distribution wherein George does have an equal (infinitesimal) probability of receiving each draft number. In this case the frequency argument for why Tracy should assign a 0.9 probability for George's demise upon learning that the game has ended stands up. In this setting, Tracy's learning that the game has ended does make an impact upon her subjective probability function, even though her prior probability for this eventuality was one. The reason is that she is already conditionalizing on George's being selected to participate in the game, and as we shall see below this event must either have probability zero, or be non-measurable.[6]

The problem is thus a special case of attempting to define a conditional probability $P(A/B)$ where B is 'measurably challenged', i.e., not of positive measure. This means that $P(B) = 0$ or that B is non-measurable. While $P(A/B)$ is typically defined as the ratio $P(A \cdot B)/P(B)$, it is usually left undefined if B is measurably challenged.[7] However, there are cases in which $P(B) = 0$, but it is intuitively clear that a conditional probability function $P(\cdot/B)$ exists.[8] (For detailed arguments on this matter, see Hájek (1999).)

As a particularly simple example, suppose that we are throwing darts at a square grid that measures one meter by one meter. Due to poor aim, the very point of the dart is equally likely to hit any point on the grid. More precisely, any two regions of equal area are equally likely to contain the point where the dart hits. Let M be the horizontal line separating the top and bottom half of the grid; in terms of the coordinate system with origin at the bottom left, M is the set of all points (x, y) with $y = 1/2$. There is zero probability that a dart will land exactly along this line. Now let L be the left half of the grid, the set of points with $x \leq \frac{1}{2}$. What is the value of $P(L/M)$, the probability that the dart lands in the region L given that it lands on the line M? Intuitively, the answer is 1/2, even though both $P(L \cdot M)$ and $P(M)$ are zero, so that the ratio $P(L \cdot M)/P(M)$ is undefined.

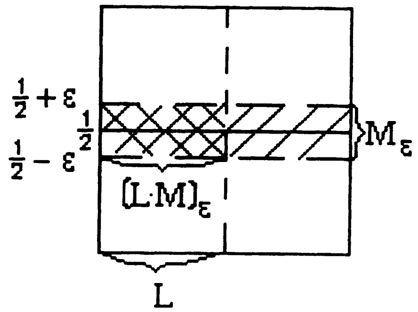


Figure 1. Conditionalizing on a set of zero measure.

Figure 1 suggests one possible justification for this claim. Let $\epsilon > 0$. Let M_ϵ be the set of all points on the grid within ϵ meters of M:

$$M_\epsilon = \{(x, y)/0 \leq x \leq 1 \text{ and } 1/2 - \epsilon < y < 1/2 + \epsilon\}.$$

Let $(L \cdot M)_\epsilon$ be the analogous extension of $L \cdot M$ by a distance of ϵ:

$$(L \cdot M)_\epsilon = \{(x, y)/(x, 1/2) \in L \cdot M \text{ and } \tfrac{1}{2} - \epsilon < y < 1/2 + \epsilon\}.$$

M can be thought of as the limit as ϵ tends to 0 of the sets M_ϵ, and analogously for $L \cdot M$, as depicted in Figure 1. Moreover, for $\epsilon > 0$, M_ϵ will have positive probability, and $P((L{\cdot}M)_\epsilon/M_\epsilon) = P((L{\cdot}M)_\epsilon{\cdot}M_\epsilon)/P(M_\epsilon)$ which is just the ratio of the areas of the two regions. But the area of M_ϵ is 2ϵ, and the area of $(L \cdot M)_\epsilon$ is 2ϵ times the linear measure of $L \cdot M$. It follows that $\lim_{\epsilon\to 0} P((L \cdot M)_\epsilon/M_\epsilon)$ is also the linear measure of $L \cdot M$. This limit seems to be an apt candidate for the conditional probability $P(L/M)$.

The above procedure can be readily generalized. Let R be any region on the grid. Then it seems intuitively clear that the conditional probability $P(R/M)$ should just be the normal linear measure (i.e., one dimensional Borel measure) of the set of points contained in the intersection of R and M, whenever this is well-defined. It suffices to expand the sets $R \cdot M$ and M to 'strips' of width 2ϵ, and to look at the limits of the conditional probabilities $P((R \cdot M)_\epsilon/M_\epsilon)$ as $\epsilon \to 0$.

The approach we take in Section 4 is modeled on this example. We will be able to define conditional probabilities $P(A/B)$, even though B is a non-measurable set (rather than a set of measure zero, as in the dartboard example). We will define $P(A/B)$ as a 'limit' of conditional probabilities of the form $P(A_\eta/B_\eta)$, where B_η is measurable and $P(B_\eta) \neq 0$. The twist is that the indices η will be non-standard numbers.

This method for ascribing conditional probabilities, when the event conditionalized on is not of positive measure, may be of much broader value in addressing problems of philosophical concern. For example, one objection to probabilistic approaches to epistemology is that they seem to endorse a certain form of dogmatism: an agent who assigns probability one to some proposition must forever after assign probability one to that proposition, no matter what else she may learn. Updating by conditionalization cannot dislodge such probability values. If, however, an agent is allowed to conditionalize on a proposition to which she did not previously assign positive probability, she may be able to retract her previous assignment of maximum probability. This is precisely what happens in the shooting room paradox. Although George and Tracy begin by assigning probability one to the proposition that the game will end, upon learning that George participates in the game, they are forced to revise their assignment to a much lower probability (5/162 to be precise!).

1.3. *Notation and Assumptions*

In order to set up the mathematics of the problem, we introduce some useful notation:

L_n	The game has length n.
R_n	George is selected to enter the room at round n of the game.
G	George is selected to enter the room at some round of the game.
N_i	George is assigned draft number i.
D	George dies.
F	The game finishes with a roll of double six. (Given the other assumptions, this means the game is finite.)
$p_i = P(N_i)$	(the prior probability for both George and Tracy that George has draft number i).
$r_n = \sum_{i=1}^{10^{n-1}} p_i$	(the prior probability that George is among the 10^{n-1} people chosen by round n). It is also appropriate to set $r_0 = 0$.

Recall that P stands for the probability function representing George and Tracy's degrees of belief (which are taken to be identical).

In addition to the assumptions enumerated in Section 1.2, we assume in what follows that the dice rolls and the assignment of draft numbers are independent, or rather that George and Tracy believe this to be the case. More precisely, this is the assumption that

$$P(N_i/L_n) = P(N_i)$$

for all i and n.[9]

1.4. *The Problem*

Recall that when George enters the shooting room, the information he learns (and on which he conditions) is just that he has been selected to enter the room, which we signify by G. Tracy, by contrast, learns that George is a participant and that the game finishes, i.e., $G \cdot F$. The two conditional probabilities of interest are thus

- $P(D/G)$ = George's posterior probability of dying, given George's information (that he has been selected to enter); and
- $P(D/G \cdot F)$ = George's posterior probability of dying, given Tracy's information (that George has been selected and the game finishes).

Assuming for the moment that these conditional probabilities are both well-defined, then the difficulty is that they should be equal; for conditionalizing on an event with probability one makes no difference, and $P(F) = 1$. Indeed, $P(F)$ is obtained by summing the probabilities that the first double-six occurs at rounds 1, 2, 3 . . . :

$$\begin{aligned}&(1/36) + (35/36)(1/36) + (35/36)^2(1/36) + \cdots \\ &= (1/36) \times \left(\frac{1}{1 - 35/36}\right) = 1.\end{aligned}$$

So it appears that the two posterior probabilities $P(D/G)$ and $P(D/G \cdot F)$ should be equal (if defined).

We divide the analysis into two cases. In the first case (Section 2), the probability function is a standard, countably additive one. In the second case (Sections 3 and 4), the probabilities are non-standard.

2. COUNTABLY ADDITIVE CASE

First suppose that the prior probability function is countably additive. Countable additivity is a standard assumption applied to probability measures. It means that if $E_1, E_2, \ldots$ are mutually exclusive (so that $P(E_i \cdot E_j) = 0$ for all $i \neq j$), then

$$P(E_1 \vee E_2 \vee \cdots) = \sum_{n=1}^{\infty} P(E_n).$$

In this case, it is false that your mother should assign probability 0.9 to your dying upon reading your name in the paper. As explained in Section

1.2, this number depends upon the assumption that George is equally likely to be assigned any draft number. The assumption is false, because the prior probabilities must converge to 0 as the draft numbers become large.

We now prove that, for this case, $P(D/G) = P(D/F \cdot G) = 1/36$ regardless of how the prior probabilities are distributed over George's positions in the draft ordering. To see this, observe that

$$(1) \qquad P(D/G) = \sum_{n=1}^{\infty} P(L_n \cdot R_n/G)$$

$$= \sum_{n=1}^{\infty} P(R_n/L_n \cdot G) P(L_n/G).$$

Now given the independence of draft numbers and dice rolls assumed above, $P(R_n/L_n \cdot G)$, the probability that George is chosen in round n given that the game has length n and George is chosen, is precisely $(r_n - r_{n-1})/r_n$. The numerator is the sum of the probabilities of nth-round draft picks; the denominator is the sum of these probabilities for all of the first n rounds. So we have

$$(2) \qquad P(R_n/L_n \cdot G) = (r_n - r_{n-1})/r_n.$$

The other term, $P(L_n/G)$, may be computed using Bayes' Theorem:

$$(3) \qquad P(L_n/G) = \frac{P(G/L_n) \cdot P(L_n)}{P(G)},$$

where (again making use of the independence of draft numbers and dice rolls)

$$P(G/L_n) = r_n;$$

$$P(L_n) = (35/36)^{n-1} \cdot 1/36;$$

$$P(G) = \sum_{n=1}^{\infty} P(G/L_n) \cdot P(L_n) = \sum_{n=1}^{\infty} r_n \cdot (35/36)^{n-1} \cdot 1/36.$$

Substituting these into (3), we get

$$(4) \qquad P(L_n/G) = \frac{(35/36)^{n-1} \cdot 1/36 \cdot r_n}{\sum_{j=1}^{\infty} r_j \cdot (35/36)^{j-1} \cdot 1/36}.$$

Notice that conditioning on G vastly increases the probability of a long game.

Substituting (2) and (4) into (1) yields

(5) $P(D/G)$

$$= \sum_{n=1}^{\infty} \frac{r_n - r_{n-1}}{r_n} \cdot \frac{r_n \cdot (35/36)^{n-1} \cdot 1/36}{\sum_{j-1}^{\infty} r_j \cdot (35/36)^{j-1} \cdot 1/36}$$

$$= \frac{1}{\sum_{j-1}^{\infty} r_j \cdot (35/36)^{j-1}} \cdot \left[\sum_{j=1}^{\infty} (r_n - r_{n-1}) \cdot (35/36)^{n-1} \right].$$

Now the expression in square brackets can be broken up into two sums, since each converges absolutely. When this is done, and indices are relabeled, the expression in square brackets becomes

$$\sum_{n=1}^{\infty} r_n \cdot (35/36)^{n-1} - (35/36) \cdot \sum_{n=0}^{\infty} r_n \cdot (35/36)^{n-1},$$

which (since $r_0 = 0$) equals

$$1/36 \cdot \sum_{n=1}^{\infty} r_n \cdot (35/36)^{n-1};$$

canceling with the first part of (5) gives the answer, 1/36.

It makes no difference if we add the information that the game is finite. This is because

$$P(F/G) = \sum_{n=1}^{\infty} P(L_n/G) = 1,$$

as is clear from an examination of (4). So if the probability measure is countably additive, there is no way to make sense of the intuition that the probability of George's having died, given that he was chosen, is 0.9.[10]

3. THE TRUNCATED SHOOTING-ROOM GAME

The analysis of Section 2 should leave us with some doubts. In particular, the requirement of a prior distribution weighted towards early draft numbers for George is disquieting. Since both George and Tracy are ignorant of the actual draft function, it seems unreasonable to weight the prior distribution in this or any other way. Why shouldn't the executioner be able to choose his victims in a random order? In any case, ignorance of the actual draft assignment should incline George and Tracy towards a symmetrical distribution which assigns equal probability to $p_i = P(N_i)$, for every i.

Suppose that such a distribution is possible, with $p_i = k$ for all i. Then Tracy might reason as follows. This time, $P(R_n/L_n \cdot G) = 0.9$, since by (2),

$$\frac{(r_n - r_{n-1})}{r_n} = \frac{(k \cdot 10^{n-1} - k \cdot 10^{n-2})}{(k \cdot 10^{n-1})} = 0.9.$$

So by (1), we have

$$(6) \qquad P(D/G) = \sum_{n=1}^{\infty} P(R_n/L_n \cdot G)P(L_n/G)$$

$$= \sum_{n-1}^{\infty} 0.9 \cdot P(L_n/G)$$

$$= (0.9) \cdot \sum_{n=1}^{\infty} P(L_n/G)$$

$$= (0.9) \cdot P(F/G),$$

which implies that $P(D/F \cdot G) = 0.9$. Let us pretend for a moment that we don't know that $P(F) = 1$. In order to reconcile the values $P(D/F \cdot G) = 0.9$ and $P(D/G) = 1/36$, we would have to have $P(F/G) = 5/162$. Given that George has been chosen, the probability that the game is infinite jumps from 0 to 157/162! This is, of course, impossible in a standard countably additive probability space; but it is precisely these values that we aim to justify.

3.1. *The Truncated Finite Game*

To motivate the approach that will be adopted, we will consider first a modified version of the problem, the "truncated finite game" of length M. As in the original version, $9 \cdot 10^{n-2}$ people are placed in the chamber on round n, and executed on a roll of double six. Unlike the original version, however, the truncated game stops even if there is no double six rolled by round M; the last group is then set free and the game is declared "safe".

Thus, we need only assume a finite population of m draftees, where $m > 10^{M-1}$. This means we may assume a uniform prior distribution: George has a $1/m$ chance of being assigned draft number i, for $1 \leq i \leq m$.

Let P stand for the associated probability measure on the possible outcomes of the game. Let T stand for "the game is safe", i.e., no double six is rolled during the first M tosses. Let G, D, and F be as before; clearly $F = \bar{T}$, since F is the event that the game finishes with a toss of double six, i.e., is unsafe. We show that $P(T/G)$ is close to 157/162: given that George enters the room, most likely it is during the last round of a safe game.

To begin, we construct an outcome space with the probability measure P. Put $\bar{m} = \{1, \ldots, m\}$, and let $\Omega_1 = \bar{m}$. Let $\Omega_2 = \mathbb{N}$. Put $\Omega_m = \Omega_1 \times \Omega_2$, and let $\mathcal{A}$ be the algebra consisting of finite unions of sets of the form $A \times B$ where $A \subseteq \Omega_1$ and $B \subseteq \Omega_2$. Events ω in Ω_m are of the form $\omega = (i, n)$ where $i \in \Omega_1$ and $n \in \Omega_2$. Here, i is George's draft position and n is the length of the game; $n > M$ represents a 'safe' game. The measure P is defined by

$$(7) \qquad P(A \times B) = \frac{|A|}{|\Omega_1|} \cdot \left[\left(\sum_{\substack{k \in B \\ k \leq M}} (35/36)^{k-1} \cdot 1/36 \right) + (35/36)^M \cdot \theta_M(B) \right],$$

where

$$\theta_M(B) = \begin{cases} 1, & \text{when } B \cap \{n/n > M\} \neq \emptyset \\ 0, & \text{otherwise.} \end{cases}$$

The probability measure reflects the independence of draft assignments and dice rolls. Note that $P(i = p) = 1/m$ for all $p \leq m$; that is, George has an equal probability of being assigned any draft number. In addition, $P(n = k) = (35/36)^{k-1} \cdot (1/36)$ if $k \leq M$, and $(35/36)^M$ if $k > M$. So P corresponds to the desired prior probability assignments.

We are interested in $P(T/G)$, the probability of a safe game given that George is picked. Figure 2 depicts the entire outcome space Ω_m.

G can be identified as the set of outcomes in the shaded region. More formally:

$$G = \{(i, n)/i \leq 10^{(M \wedge n)-1}\},$$

where $M \wedge n$ is the minimum of M and n. For we need both $i \leq 10^{n-1}$ and $i \leq 10^{M-1}$ in order to have $(i, n) \in G$. The set T of safe games is just the set of outcomes where $n > M$:

$$T = \{(i, n)/n > M\}.$$

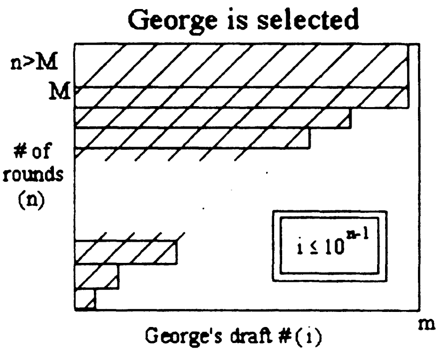


Figure 2. The truncated finite game.

To compute $P(T/G) = P(T \cap G)/P(G)$, we need to compute $P(T \cap G)$ and $P(G)$.

$$(8) \qquad P(G) = \sum_{j=1}^{M} P(i \leq 10^{j-1}, n = j) + P(i \leq 10^{M-1}, n > M)$$

$$= \sum_{j=1}^{M} \frac{10^{j-1}}{m}(35/36)^{j-1}(1/36) + \frac{10^{M-1}}{m}(35/36)^{M}$$

$$= \frac{1}{36m}\frac{(350/36)^{M} - 1}{(350/36) - 1} + \frac{1}{10m}\left(\frac{350}{36}\right)^{M}$$

$$= \frac{1}{314m}[(350/36)^{M} - 1] + \frac{1}{10m}\left(\frac{350}{36}\right)^{M}.$$

$$(9) \qquad P(T \cap G) = \frac{10^{M-1}}{m}(35/36)^{M} = \frac{1}{10m}(350/36)^{M}.$$

So dividing (9) by (8), we obtain

$$(10) \qquad P(T/G) = \frac{1}{\frac{10}{314}[1 - (\frac{36}{350})^{M}] + 1}$$

$$= \frac{1}{\frac{162}{157} - \frac{10}{314}(\frac{36}{350})^{M}}.$$

Notice that this number is independent of the population size m, and clearly converges to 157/162 as $M \to \infty$.

This means that if F stands for "the game ends with double six", we have $P(F/G) \approx 5/162$. Notice how different this is from the prior probability $P(F) = 1 - (35/36)^M$, which converges to 1 as $M \to \infty$. The most important observation here is this: although a safe game is *a priori* unlikely, once we conditionalize on the vastly more unlikely event that George is picked, the safe game becomes extremely probable.

4. A NONSTANDARD MODEL OF THE SHOOTING-ROOM GAME

In this section we will use nonstandard analysis to construct a new model of the shooting-room game. We stress that while we will be using nonstandard analysis as a tool, the probability measures that we ultimately define will be strictly real-valued, and finitely additive. Thus we are not committed to the existence of infinitesimal degrees of belief or anything of that sort. Just as imaginary numbers can be used to facilitate the proving of theorems that exclusively concern real numbers, our use of nonstandard analysis will be used to facilitate and motivate the construction of purely real-valued measures.

In Section 4.1, we translate the truncated finite shooting-room game, described in the last section, into the milieu of non-standard analysis. This will yield a model of the game that we call the "truncated hyperfinite game", in which both the potential number of rounds and the population of people eligible to participate in the game are "hyperfinite" or nonstandardly infinite. Many readers will no doubt feel some trepidation upon seeing these words. The foundations of nonstandard analysis are indeed mathematically complex (as are the foundations of real analysis, we hasten to add), but once one has the apparatus at hand, it is actually quite easy to function within the milieu of nonstandard analysis. In particular, the infinite and infinitesimal numbers that populate the nonstandard realm are subject to the rules of finitistic mathematics in a way that the more familiar infinities of Cantorian set theory are not. Any readers able to comprehend the truncated finite model of the previous section should be able to read Section 4.1 with little difficulty.

The fact that the infinities that appear in nonstandard models are quite different from those that we encounter in set theory does, however, give rise to a complication: the model we construct in Section 4.1 does not quite fit the shooting-room game as originally described. In particular, the latter assumed a countably infinite population, while our model treats the population as "hyperfinite". The model of 4.1 allows, for example, that the game might go on for some infinite number of rounds with no double sixes being rolled, and then stop after some specific round. This possibility of stopping *after* some infinite number of rounds seems alien to our normal way of thinking about infinity. We do not believe that this would be too high a price to pay for a solution to the shooting-room paradox, but fortunately, we do not have to pay it. In Sections 4.2 and 4.3, we gradually convert the "truncated hyperfinite game" of Section 4.1 back into a standard model of the game, where the number of potential rounds and of potential participants is countably infinite. On our way there, we will construct a "semi-standard" game where the population is hyperfinite, but the number of rounds in the game is standard, i.e., potentially countably infinite. Sections 4.2 and 4.3 provide only an overview of this construction; the mathematical details are relegated to Section 6. This material is somewhat more technical, and those readers who do not wish to bother themselves with the details of converting non-standard infinities to standard infinities may skip Sections 4.2 and 4.3 with little loss of continuity.

We begin with a very brief review of nonstandard analysis. The central concept is the *-transform (pronounced star-transform). This is a function that maps standard entities into their nonstandard counterparts. The domain of this function is the set-theoretic hierarchy erected upon the real numbers. This domain includes real numbers, sets of real numbers, relations, functions, sequences, and so on. The image of some standard entity s under the *-transform will be denoted $^*s$. For example, the set of standard numbers $\mathbb{N}$ becomes the nonstandard set $^*\mathbb{N}$. This entity will also be a set; it will not, however, just be the set of images of natural numbers under the *-transform. (We will elaborate upon this fact momentarily.) Indeed, this latter set $\{^*n/n \in \mathbb{N}\}$ (which it is convenient to refer to as $\mathbb{N}$, even though it is not, strictly, a standard entity) is not the *-transform of any standard entity.

This latter observation points to an important distinction between *internal* and *external* nonstandard entities. The definition of internal entities is complex, but roughly speaking, an entity is internal if it belongs to a set which is the image of some standard entity under the *-transform; otherwise, it is *external*.[11] $\mathbb{N}$ is an external set; intuitively, from the perspective of nonstandard analysis, this is not a natural set of numbers, but more of a set-theoretic gerrymander. In general, Cantorian infinities are very different from nonstandard infinities, and the two mesh together very awkwardly.

For illustrative purposes, we give an oversimplified and unrigorous analogy. (Those curious about the rigorous details should consult Hurd and Loeb 1985, chapters 1 and 2.) Let * map each real number onto an equivalence class of sequences of real numbers. $^*n$ will be the equivalence class of $\langle n, n, \ldots \rangle$, or $[\langle n, n, \ldots \rangle]$. Two sequences will be equivalent if their terms are identical in 'enough' positions. We will not say precisely what 'enough' means, but if two sequences agree in all but finitely many places, they agree in 'enough' places.

More generally, the *-transform of a standard number-theoretic property will hold of a nonstandard number if, for one of its representative sequences, that property holds for 'enough' standard numbers in the sequence. Consider, for example, the sequence $\langle 1, 2, 3, 4, \ldots \rangle$. Since each term in this sequence belongs to $\mathbb{N}$, $[\langle 1, 2, 3, \ldots \rangle]$ will belong to $^*\mathbb{N}$. Note, however, that $[\langle 1, 2, 3, \ldots \rangle]$ is not equal to $^*n$ for any standard natural number n. Indeed, $[\langle 1, 2, 3, \ldots \rangle]$ is larger than any such number, since all but finitely many of the terms of the sequence are larger than n. Thus $[\langle 1, 2, 3, \ldots \rangle]$ is infinite, or more properly, *hyperfinite*. To see that such hyperfinite numbers are different from Cantorian infinite numbers, the reader should convince herself that there is no smallest hyperfinite number. Analogously, the number $[\langle 1, 1/2, 1/3, \ldots \rangle]$, is greater than zero, yet smaller than any real number. Such numbers are *infinitesimal*. The *-transforms of the normal mathematical operations of addition, subtraction, multiplication and division are all well-defined on hyperfinite and infinitesimal numbers. The reader may confirm, for instance that the product of these two nonstandard numbers is ^{*1}.

Finally, we state two very useful facts. First, any finite non-standard real number can be decomposed into the sum of a finite real number, known as its *standard part*, and an infinitesimal number.[12] We write $^\circ\eta$ for the standard part of η. Second, let $s = s_1, s_2, \ldots$ be a sequence of standard numbers that converges to limit n. Formally, s is a function defined on the natural numbers. Then $^*s$ will be a function on $^*\mathbb{N}$ such that for every hyperfinite argument η, s_η differs from n by at most an infinitesimal. In other words, a standard sequence converges to a value if and only if its nonstandard counterpart gets (and stays) infinitely close to that value.

4.1. *The Truncated Hyperfinite Game*

We first construct a non-standard version of the truncated game. Instead of stopping the game after a finite number of rounds, however, we stop if no double six occurs in the first η rounds, where η is an infinite integer in $^*\mathbb{N}\backslash\mathbb{N}$. Instead of a countable population, we now need to assume a hyperfinite population of size m, for some fixed $m \epsilon {}^*\mathbb{N}$, $m \geq 10^{\eta-1}$. This allows us to define a uniform probability distribution for the draft assignments. As in Section 3.1, we will construct a (non-standard) probability measure such that George has equal probability $1/m$ of being assigned any place in the draft.

The outcome space is $\Omega_m = \Omega_1 \times \Omega_2$, where $\Omega_1 = \bar{m}$ and $\Omega_2 = {}^*\mathbb{N}$.[13] The algebra $\mathcal{A}$ consists of hyperfinite unions of sets $A \times B$ where A is an internal subset of Ω_1 and B is an internal subset of Ω_2.

Again, we write $\omega = (i, n)$ for elements of Ω_m, where $i \in \Omega_1$ and $n \in \Omega_2$. A safe game is represented by the condition $n > \eta$. We define a non-standard measure μ_η on the algebra $\mathcal{A}$ by analogy with the definition for the truncated finite game

$$(11) \quad \mu_\eta(A \times B) = \frac{|A|}{|\Omega_1|} \cdot \left[\left(\sum_{\substack{k \in B \\ k \leq \eta}} (35/36)^{k-1} \cdot 1/36 \right) + (35/36)^\eta \cdot \theta_\eta(B) \right],$$

where $|A|$ is the internal cardinality of A and

$$\theta_\eta(B) = \begin{cases} 1, & \text{when } B \cap \{n/n > \eta\} \neq \emptyset \\ 0, & \text{otherwise.} \end{cases}$$

Then $(\Omega_m, \mathcal{A}, \mu_\eta)$ is the internal probability space for the truncated game of length η. Although the measure μ_η depends on the choice of η, we shall refer to it as μ for the remainder of this section, since η is held constant. As before, $\mu(\{(i, n)/i = p\}) = 1/m$ for all $p \leq m$, so μ gives us the required uniform distribution over draft positions.

Once again, we are interested in obtaining $P(T/G)$, George's and Tracy's subjective probability of a safe game, given that George is picked. This value can be determined using the non-standard probability μ. Here,

$$G = \{(i, n)/i \leq 10^{(\eta \wedge n)-1}\}$$

is the set of all outcomes in which George is picked, where $\eta \wedge n$ is the minimum of η and n. Also,

$$T = \{(i, n)/n > \eta\}$$

is the set of 'safe' games. These sets are analogous to their counterparts in the truncated finite game of Section 3.1. Figure 3 provides a picture of the probability space Ω_m, letting the shaded region represent the outcomes in the set G.

The calculations (8)–(10) proceed exactly as in the finite case, replacing M with η. In particular, if we write μ_G for the conditional probability $\mu(\cdot/G)$, then μ_G is a non-standard measure on $\mathcal{A}$, and we have

$$(12) \quad \mu(G) = \frac{1}{314m}[(350/36)^\eta - 1] + \frac{1}{10m}\left(\frac{350}{36}\right)^\eta$$

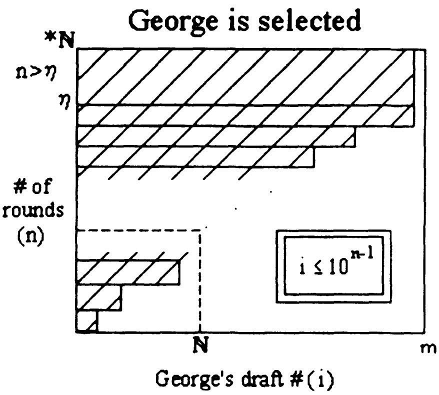


Figure 3. The truncated hyperfinite game.

and

$$(13) \qquad \mu_G(T) = \frac{1}{\frac{162}{157} - \frac{10}{314}(\frac{36}{350})^\eta}.$$

Thus, $\mu_G(T)$ differs from 157/162 by an infinitesimal amount.

Now Loeb (1975) has shown how to turn a non-standard measure μ on an internal algebra $\mathcal{A}$ into a standard (real-valued, countably additive) measure $\hat{\mu}$ on the smallest σ-algebra containing $\mathcal{A}$. The Loeb construction involves putting $\hat{\mu}(A) = {}^\circ\mu(A)$ for sets $A \in \mathcal{A}$, where ${}^\circ\mu(A)$ is the standard part of the number $\mu(A)$, and then utilizing Carathéodory's technique for extending a finitely additive measure on an algebra to a countably additive measure on a σ-algebra. Applying the Loeb construction to μ_G instead of μ, we obtain a standard, real-valued measure $\hat{\mu}_G$ such that, $\hat{\mu}_G(T) = 157/162$. The measure $\hat{\mu}_G$ is precisely what we mean by $P(\cdot/G)$, the probability conditional on G. If, as before, we put $F = \bar{T}$, then F represents an unsafe game in which a double six occurs by round η. Then $\hat{\mu}_G(F) = 5/162$.

Calculations similar to (8)–(10) show that $\hat{\mu}(D) = 1/36$, i.e., the chance of George's dying (or having died) given that he is chosen is 1/36, as it should be. Here,

$$D \cap G = \{(i, n)/10^{n-2} < i \leq 10^{n-1} \text{ and } n \leq \eta\},$$

so

$$(14) \qquad \mu(D \cap G) = \sum_{j=1}^{\eta} \frac{10^{j-1} - 10^{j-2}}{m} \left(\frac{35}{36}\right)^{j-1} \frac{1}{36}$$

$$= \frac{0.9}{36m} \frac{(350/36)^{\eta} - 1}{(350/36) - 1}$$

$$= \frac{0.9}{314m}[(350/36)^{\eta} - 1].$$

Dividing (14) by $\mu(G)$ as given in (12), we obtain

$$\frac{0.9[1 - (36/350)^{\eta}]}{(324/10) - (36/350)^{\eta}}$$

which has standard part 1/36.

4.2. *The Semi-standard Game*

So far, we have shown that in the two truncated versions of the game, George's and Tracy's inferences are both correct. This is because in each of these cases, we have $P(D/G) = 1/36$, $P(D/F \cdot G) = 0.9$ and $P(F/G) = \frac{5}{162}$. So the probability that George dies given that he has been chosen is 1/36, but once we condition on the further information that the game is unsafe, this probability rises to 0.9. This accomplishes one of our objectives: showing how $P(D/G)$ and $P(D/F \cdot G)$ can be different, given a uniform distribution.

But it is dissatisfying to end here. While the truncated hyperfinite game yields the correct probability values, it does so by introducing into the outcome space bizarre events which we had not originally envisaged, such as George's receiving a hyperfinite draft number, or participating in the game during a hyperfinite round. We will eventually construct a model which yields the right probabilities, but on a completely standard outcome space in which there is neither a hyperfinite population nor hyperfinite rounds. The construction proceeds in two stages. First, in this section, we consider a 'semi-standard' version of the game which involves a hyperfinite population, but no hyperfinite rounds. In the next section, we solve the original problem by restoring the assumption of a countable population.

Caveat: this section and the next involve a marked increase in technical difficulty. The move to a countable population and at most countably many rounds results in some of the relevant events (George's being picked; the game's being finite) becoming non-measurable. This complicates matters significantly. Readers who are satisfied with the solution provided thus far may wish to skip to Section 5.

The outcome space for the semi-standard game is the same as in section 4.1: $\Omega_m = \bar{m} \times {}^*\mathbb{N}$, where we fix a hyperfinite integer m, the size of the population. As before, let $\bar{m} = \{1, 2, \ldots, m\}$. Then $\Omega = \{(i, n)/i \in \bar{m}$

and $n \in {}^*\mathbb{N}\}$, where as usual i is George's draft number and n is the length of the game. The games that don't finish are just the non-finite outcomes, that is the pairs (i, n) with $n \notin \mathbb{N}$. The unsafe games are thus the finite outcomes:

$$F = \{(i, n)/n \in \mathbb{N}\}.$$

G is the set of games in which George is chosen at some finite round during the game:

$$G = \{(i, n)/i \in \mathbb{N}, i \leq 10^{n-1}\}.$$

Although i must be finite for George to be chosen, n need not be in $\mathbb{N}$, since the set G includes infinite games in which George is picked.

We define a measure μ on the algebra $\mathcal{A}$ of internal subsets of Ω_m by setting

$$\text{(15)} \qquad \mu(A \times B) = \frac{|A|}{|\bar{m}|} \cdot \left(\sum_{k \in B} (35/36)^{k-1} \cdot 1/36 \right).$$

The hyperfinite sum is well-defined, because B is internal. μ is defined on any hyperfinite union of rectangles $A \times B$. So $(\Omega_m, \mathcal{A}, \mu)$ is an internal probability space. Moreover, $\mu(\{(i, n)/i = p\}) = 1/m$ for all $p \in \mathbb{N}$, so μ gives the required uniform distribution over draft positions. Note that μ and the μ_η of Section 4.1 (defined in (11)) are very similar in form, but μ does not depend upon η.

Since (as noted earlier) $\mathbb{N}$ is not internal, neither F nor G is internal; hence, neither set is μ-measurable. However, both F and G can be written as countable unions of sets in $\mathcal{A}$. Consequently, if $\hat{\mu}$ is obtained from μ by the Loeb construction, then $\hat{\mu}(G)$ and $\hat{\mu}(F)$ are defined, and in fact $\hat{\mu}(G) = 0$ and $\hat{\mu}(F) = 1$.[14] So it is natural to think of G as having probability 0 and F as having probability 1. Of course, the fact that $\hat{\mu}(G) = 0$ implies that we cannot define conditional probabilities $\hat{\mu}(\cdot/G)$ in the usual way.

Nevertheless, we can imagine conditionalizing on the event that George is picked. We can, for instance, feel fairly confident that $P(D/G)$ is 1/36. We might also be led by our experience with the two truncated games to believe that $P(D/G \cdot F) = 0.9$. If we are prepared to go this far, then we must also believe that $P(F/G) = 5/162$. There might also be additional events E such that $P(E/G)$ has an intuitive value.

But how can we make sense of conditionalizing on a set that is either non-measurable or has measure zero? The most convincing answer is to demonstrate that we can actually construct a measure μ_G on an algebra $\mathcal{A}_{\mathcal{G}}$ of subsets of Ω_m that has the required properties. The remainder of this section accomplishes this task; we stress that the construction here is the key to the solution of the original problem.

We construct the algebra $\mathcal{A}_{\mathcal{G}}$ in such a way that for $\eta \in {}^*\mathbb{N}\backslash\mathbb{N}$, every A in $\mathcal{A}_{\mathcal{G}}$ has a natural analogue A_η in the algebra for the truncated hyperfinite game of length η. In fact, each A will be the 'setwise limit' of the sets A_η, since the intersection over $\eta \in {}^*\mathbb{N}\backslash\mathbb{N}$ of the set theoretic differences $A \triangle A_\eta$ is empty. This means that there is nothing included in (excluded from) A that is not eventually included in (excluded from) all sets A_η, as η decreases within ${}^*\mathbb{N}\backslash\mathbb{N}$.

Even though some of the sets in $\mathcal{A}_{\mathcal{G}}$ will not be internal, each set A_η will be an internal subset of Ω_m. We show that the standard part of the quotient

$$(16) \qquad \frac{\mu(A_\eta \cap G_\eta)}{\mu(G_\eta)}$$

is constant for each η, and define a conditional probability measure μ_G on $\mathcal{A}_{\mathcal{G}}$ by letting $\mu_G(A)$ be this constant. This process gives us a well-defined, finitely additive probability measure on an algebra of sets.[15]

This strategy for defining the conditional probability is similar in spirit to the method described for the geometric example in the introduction. We expand both A and G slightly to sets with a well-defined, non-zero μ-measure, and take the standard parts of the quotients in (16). It turns out that these standard parts are constant, which is the analogue of convergence in the truncated finite game.

The algebra $\mathcal{A}_{\mathcal{G}}$ and the η-mapping are constructed in Section 6.1; the measure μ_G is constructed and proven to be well-defined in Section 6.2. We will here only summarize the properties of μ_G:

- The measure μ_G is finitely additive on the algebra $\mathcal{A}_{\mathcal{G}}$.
- $\mu_G(F) = \frac{5}{162}$ which follows from the result (13) of Section 4.1. The chance of a finite game is slim, given that George is picked.
- $\mu_G(D) = \frac{1}{36}$, which follows from the discussion following (14) of Section 4.1.
- $\mu_G(D/F) = 0.9$, which follows from the above results and the fact that $\mu_G(D/\bar{F}) = 0$.

4.3. *The Original Game*

At last! We are finally ready to construct a measure ν_G which intuitively corresponds to conditionalizing on George's being picked in the original game, i.e., to $P(\cdot/G)$. Recall that the original game means: no hyperfinite population and no hyperfinite rounds of the game. However, we still need to use hyperfinite integers to represent the game, in order to establish a uniform probability distribution and to represent the possibility of an infinite game.

The outcome space $\Omega = \mathbb{N} \times {}^*\mathbb{N}$. So $\Omega = \{(i, n)/i \in \mathbb{N} \text{ and } n \in {}^*\mathbb{N}\}$, where i is George's draft number and n is the length of the game. As in Section 4.1, the unsafe games are simply the finite outcomes:

$$F = \{(i, n)/n \in \mathbb{N}\}.$$

And G is the set of outcomes in which George is chosen at some finite round during the game (we no longer need to explicitly require $i \in \mathbb{N}$):

$$G = \{(i, n)/i \leq 10^{n-1}\}.$$

It is straightforward to turn this set into a probability space with the correct unconditional probabilities for events such as "the game lasts n rounds", or "George is picked on round n". The events are those in the algebra $\mathcal{A}$ consisting of finite unions of sets $A \times B$ where A is any subset of $\mathbb{N}$ and B is an internal subset of ${}^*\mathbb{N}$. We want a measure which assigns an equal infinitesimal probability to George's being given any draft number. The measure on $\mathcal{A}$, which we shall call ν, is defined as follows. Fix a hyper-finite integer m, and as before let $\bar{m} = \{1, 2, \ldots, m\}$. This set has a non-standard cardinality of m. For any $A \subseteq \mathbb{N}$ and internal $B \subseteq {}^*\mathbb{N}$, set

$$(17) \qquad \nu(A \times B) = \frac{|{}^*A \cap \bar{m}|}{|\bar{m}|} \cdot \left(\sum_{k \in B} (35/36)^{k-1} \cdot 1/36 \right),$$

where $|{}^*A \cap \bar{m}|$ is the internal cardinality of ${}^*A \cap \bar{m}$.[16] The hyperfinite sum is well-defined (since B is internal). So $(\Omega, \mathcal{A}, \nu)$ is a probability space. Most importantly, $\nu(\{(i, n)/i = p\}) = 1/m$ for all $i \in \mathbb{N}$; thus ν gives the required uniform distribution by assigning equal probability to each of the countably many draft positions.[17] In addition, ν assigns probability $(35/36)^{k-1} \cdot 1/36$ to the event $n = k$. So ν is just what we want to represent George and Tracy's prior probabilities for the shooting-room game. And nothing in the definition of ν depends on the preceding sections.

The trouble arises, of course, when we try to conditionalize on G, because neither G nor F is ν-measurable. However, we can conditionalize on G by simply carrying over the work of Section 4.2, once we observe a systematic correspondence between the probability spaces defined in this section and the last.

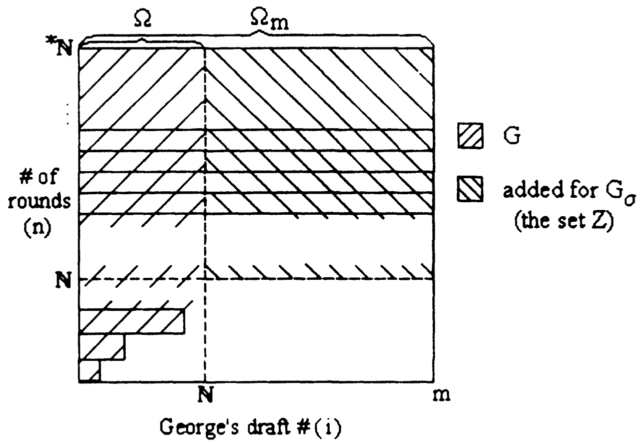


Figure 4. The σ-mapping.

Note first that for any rectangle $A \times B$ in Ω, the rectangle $({}^*A \cap \bar{m}) \times B$ is a measurable subset of the outcome space $\Omega_m = \bar{m} \times {}^*\mathbb{N}$ defined in Section 4.2, since ${}^*A \cap \bar{m}$ is internal. Furthermore, it is clear that

$$\nu(A \times B) = \mu(({}^*A \cap \bar{m}) \times B).$$

This correspondence defines a mapping between standard outcomes in Ω and 'semi-standard' outcomes in Ω_m; we shall call it the σ-mapping. Figure 4 illustrates the correspondence. In the lower part of the picture, where A is a finite subset of $\mathbb{N}$, horizontal rectangles of the form $A \times B$ map to themselves, since ${}^*A = A$. The rectangles above the dotted line (signifying $n \notin \mathbb{N}$) are all of the form $\mathbb{N} \times \{n\}$ and are mapped to $\bar{m} \times \{n\}$.

We want to extend the σ-mapping to sets such as G, as illustrated. In order to do this, we interpret the σ-mapping in the following manner. The way that we have defined the measure ν on Ω is to regard a standard event E, definable as a ν-measurable subset of Ω, as analogous to a μ-measurable subset of Ω_M. This latter subset can be thought of as representing the analogous event $\sigma(E)$ in the 'semi-standard' game: it is the same event, except that a hyperfinite population of size m is substituted for the countable population of the standard set-up. We then define the probability of the standard event E by assigning to it the probability of its 'semi-standard' analogue $\sigma(E)$.

For example, consider the event, "the game lasts more than 1 round". This can be represented as $J = \{(i, n)/n > 1, i \in \mathbb{N}\}$ for the standard game, and $\sigma(J) = \{(i, n)/n > 1, i \in \bar{m}\}$ for the semi-standard game. Of course, $\nu(J) = \mu(\sigma(J))$. Similarly, consider the event, "George's draft number is greater than 10". The standard representation is $K = \{(i, n)/i >$

10, $i \in \mathbb{N}\}$; the semi-standard representation is $\sigma(K) = \{(i,n)/i > 10, i \in \bar{m}\}$, and once again $\nu(K) = \mu(\sigma(K))$.

The crucial next step is to extend the mapping σ between standard and semi-standard events to non-measurable events – specifically, to those that can be represented in the algebra $\mathcal{A}_{\mathcal{G}}$ of Section 4.2. Most importantly we want to establish analogues for the two events of greatest interest in the shooting-room game: F (the game finishes) and G (George is picked). The analysis is clearest for F. The 'unsafe' standard game

$$F = \{(i,n)/n \in \mathbb{N}, i \in \mathbb{N}\}$$

plainly corresponds to the unsafe semi-standard game

$$\sigma(F) = \{(i,n)/n \in \mathbb{N}, i \in \bar{m}\}.$$

More problematically, the event of George's being picked in the standard game,

$$G = \{(i,n)/i \in \mathbb{N} \quad \text{and} \quad i \leq 10^{n-1}\},$$

corresponds to George's being picked in the non-standard game, which happens to be *exactly* the same set:

$$\sigma(G) = \{(i,n)/i \in \mathbb{N} \quad \text{and} \quad i \leq 10^{n-1}\}.$$

For even in the semi-standard game, George's draft number must be in $\mathbb{N}$ if he is actually selected. Arguably, however, $\sigma(G)$ should be the larger set pictured in Figure 4, since this is what we would get if we applied the σ-mapping one rectangle at a time.

This creates a technical difficulty, because we want to define σ on a basis for an algebra of sets, and extend it inductively to the full algebra by putting $\sigma(A \cap B) = \sigma(A) \cap \sigma(B)$ and $\sigma(A \cup B) = \sigma(A) \cup \sigma(B)$. In precise terms, the difficulty is as follows. Consider the standard set $\bar{G} \cap \bar{F}$ of infinite games in which George is not selected. In the standard game, this set is empty: George is assigned a finite draft number, so that he will eventually be selected in any infinite game. By contrast, in the semi-standard game, this set is not empty, since there are infinite games in which George is assigned a draft number in $\bar{m} \backslash \mathbb{N}$.

In fact, if we let $\sigma(\bar{G})$ and $\sigma(\bar{F})$ stand for the semi-standard outcomes, then

$$\sigma(\bar{G}) \cap \sigma(\bar{F}) = \{(i,n)/i \in \bar{m} \backslash \mathbb{N}, n \in {}^*\mathbb{N} \backslash \mathbb{N}\};$$

we will call this set Z. The set Z is indicated in Figure 4. The difficulty, then, is that the mapping σ is not well-defined: the image $\sigma(\emptyset)$ of the empty set $\emptyset$ should again be the empty set, but by writing $\emptyset = \bar{G} \cap \bar{F}$ and applying the inductive definition, we get $\sigma(\emptyset) = Z$.

Fortunately, it is not too difficult to tidy up σ, because it turns out that $\mu_G(Z) = 0$, where μ_G is the conditional probability defined in Section 4.2. It should hardly be surprising that the probability of an infinite (semi-standard) game in which George is *not selected*, given that George *is selected*, is zero, and the proof is straightforward. Because Z is a set of measure 0, we can simply tack it on in defining the σ-image. This will allow us to define the conditional probability ν_G for the standard game, on essentially the same algebra of sets as in Section 4.2. The details are contained in Section 6.3.

We may thus define, for A in the algebra $\mathcal{A}_{\mathcal{G}}$,

$$\nu_G(A) = \mu_G(\sigma(A)).$$ [18]

This completes the construction of the probability measure ν_G, which is intended to have the properties of $P(\cdot/G)$, probability conditionalized on G. The measure is finitely additive on the algebra $\mathcal{A}_{\mathcal{G}}$.

It is worth briefly rehearsing the steps involved in evaluating $\nu_G(A)$, for $A \in \mathcal{A}_{\mathcal{G}}$.

Step 1: The σ-mapping. First, we map A, a subset of $\Omega = \mathbb{N} \times {}^*\mathbb{N}$, to $\sigma(A)$, the analogous subset of $\Omega_m = \bar{m} \times {}^*\mathbb{N}$. This sets up a correspondence between events in the standard game and events in the semi-standard game; the measure ν on Ω, corresponds to μ on Ω_m.

Step 2: The η-mapping. Second, we modify both $\sigma(A)$ and G – neither of which is likely to be η-measurable – by applying the η-mapping to both sets (for suitably small $\eta \in {}^*\mathbb{N}\backslash\mathbb{N}$). In effect, this sets up a correspondence between the semi-standard game and the truncated hyperfinite game of Section 4.1. Theorem 1 assures us that the ratios $\mu(\sigma(A)_\eta \cap G_\eta)/\mu(G_\eta)$ are constant (up to an infinitesimal difference).

Step 3: Conditionalization. $\nu_G(A)$ is defined as this constant ratio.

Since $\mu_G(Z) = 0$, we may carry over all important results from Section 4.2. In particular, we have at last proved the results, which all follow from the discussion at the end of Section 4.2:

- $\nu_G(F) = \frac{5}{162}$.
- $\nu_G(D) = \frac{1}{36}$.

- $\nu_G(D/F) = 0.9$.

5. CONCLUSION

In the countably additive case, there is no way to make sense of assigning a subjective probability of 0.9 to George's demise – but the weakness of this analysis is its inability to accommodate the intuition that George is equally likely to have any draft position. In the non-standard case, which allows us to assign equal probability to each event in a countable partition, the above analysis demonstrates that the crucial factor in determining whether the subjective probability of George's dying is 1/36 or 0.9 is the possession of knowledge about whether or not the game comes to an end.[19]

This only confirms what we all know: mother is always right, or at least never wrong! George, like so many of his generation, was quick to accuse his mother of worrying too much.[20] But with experience (and a little nonstandard analysis), George will soon come to realize that his mother's fear for his life was justified after all.

6. CONSTRUCTIONS AND PROOFS

6.1. *Construction of $\mathcal{A}_{\mathcal{G}}$ and the η-mapping*

In this section, we construct the family $\mathcal{A}_{\mathcal{G}}$ and the η-mapping of Section 4.2.

In what follows, we use ϕ to represent an arbitrary function from $\mathbb{N}$ to $\mathbb{N}$, and ψ for the specific function $\psi(n) = 10^{n-1}$. For any such ϕ, put $\phi'(n) = \phi(n) \wedge \psi(n)$, the minimum of ϕ and ψ. We shall say that ϕ is *of the same order as* ψ if $\lim_{n\to\infty}\phi(n)/\psi(n)$ exists; the limit is a non-negative real number. This condition will be abbreviated by writing ϕ is $\bigcirc\psi$; we will also write

$$\frac{\phi(n)}{\psi(n)} \sim k$$

to mean that $\lim_{n\to\infty}\phi(n)/\psi(n) = k$. Note that if ϕ' is $\bigcirc\psi$, then $\phi'(n)/\psi(n) \sim k$ for some k with $0 \leq k \leq 1$.

We start with a *basis* β for the algebra, consisting of four types of subsets of Ω_m.

1. $A_\phi = \{(i, n)/i \leq \phi(n) \text{ and } i \in \mathbb{N}\}$, where ϕ' is $\bigcirc\psi$.
2. $\bar{A}_\phi = \{(i, n)/i > \phi(n) \text{ or } i \notin \mathbb{N}\}$, where ϕ' is $\bigcirc\psi$.

3. $B_S = \{(i,n)/n \in S\}$, where $S \subseteq \mathbb{N}$ is finite or co-finite.
4. $\bar{B}_S = \{(i,n)/n \notin S\}$, where $S \subseteq \mathbb{N}$ is finite or co-finite.

The algebra $\mathcal{A}_{\mathcal{G}}$ consists of all finite unions of finite intersections of sets in β. It is easy to verify that this set is closed under finite intersections, finite unions, and the taking of complements; hence, it is an algebra. Most standard events of interest in the game can be expressed in terms of a set in this algebra. Sets of types 3 and 4 represent conditions on the length of the game, while in sets of type 1 and 2, the functions ϕ are used to specify ranges of possible draft positions for George. For instance, $10^{n-2} < i \leq 10^{n-1}$ represents the condition that George is chosen in the last round; the set of outcomes satisfying this condition is $\bar{A}_\phi \cap A_\psi$, where $\phi(n) = 10^{n-2}$.

It is easy to check that the intersection of any two sets of the same type is another set of that type. This is obvious for sets of type 3 or 4, and for the first two types it is a consequence of the following Lemma.

LEMMA 1. If ϕ_1 and ϕ_2 are $\bigcirc\psi$, then so are $\phi_1 + \phi_2$, $|\phi_1 - \phi_2|$, $\phi_1 \wedge \phi_2$, $\phi_1 \vee \phi_2$ and $k \cdot \phi_1$, for any constant $k \in \mathbb{N}$. Here, $\wedge$ denotes minimum and $\vee$ denotes maximum.

Proof. Suppose $\phi_1(m)/\psi(n) \sim a$ and $\phi_2(n)/\psi(n) \sim b$. Then we can write $\phi_1(n) = a \cdot \psi(n) + \alpha(n)$ and $\phi_2(n) = b \cdot \psi(n) + \beta(n)$, where $\alpha(n)/\psi(n) \sim 0$ and $\beta(n)/\psi(n) \sim 0$. Since $[\alpha(n) + \beta(n)]/\psi(n)$ clearly approaches 0 as $n \to \infty$, this shows that $[\phi_1(n) + \phi_2(n)]/\psi(n) \sim a + b$. The other results are proved similarly. □

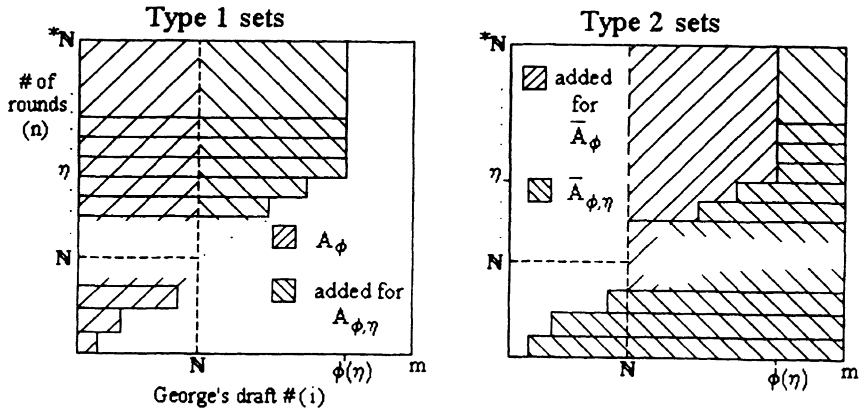
For sets of types 1 and 2, we allow only functions ϕ such that ϕ' is $\bigcirc\psi$. This is because $\phi'/\psi \sim k$ if and only if $^\circ[\phi'(\eta)/\psi(\eta)] = k$ for every $\eta \in {}^*\mathbb{N}\backslash\mathbb{N}$,[21] and it can be verified that ϕ' must satisfy this condition in order for $\mu_G(A_\phi)$ to be defined.

The next step is to define, for $\eta \in {}^*\mathbb{N}\backslash\mathbb{N}$, the η-mapping which pairs each set A in $\mathcal{A}_{\mathcal{G}}$ with an internal subset A_η of Ω_m. We first define A_η for the four basic types of sets.

1. $A_{\phi,\eta} = \{(i,n)/i \leq \phi(n \wedge \eta)\}$.
2. $\bar{A}_{\phi,\eta} = \{(i,n)/i > \phi(n \wedge \eta)\}$.
3. $B_{S,\eta} = \{(i,n)/n \in {}^*S$ and $n \leq \eta\}$.[22]
4. $\bar{B}_{S,\eta} = \{(i,n)/n \notin {}^*S$ or $n > \eta\}$.

The following pictures illustrate the relation between sets of types 1 and 2 in $\mathcal{A}_{\mathcal{G}}$ and their images under the η-mapping; similar relations exist for types 3 and 4. The pictures make it clear that for any $A \in \mathcal{A}_{\mathcal{G}}$, the sets A_η 'converge setwise' to A as η decreases within ${}^*\mathbb{N}\backslash\mathbb{N}$.

Strictly speaking, for unbounded functions ϕ, the sets $A_{\phi,\eta}$ and $\bar{A}_{\phi,\eta}$ are not proper subsets of Ω_m if η is sufficiently large, since there will be



[21] ϕ and ψ are extended to the hyperfinite integers via the *-transform.
[22] By $^*S$, we mean the *-transform of S.

Figure 5. The η-mapping.

values of η with $\phi(\eta) > m$. It is sufficient for our purposes that given $m \in {}^*\mathbb{N}\backslash\mathbb{N}$ and given any function ϕ, there is a $\Lambda \in {}^*\mathbb{N}\backslash\mathbb{N}$ such that if $\eta \leq \Lambda$, then $\phi(\eta) \leq m$. In other words, for sufficiently small $\eta \in {}^*\mathbb{N}\backslash\mathbb{N}$, the sets $A_{\phi,\eta}$ and $\bar{A}_{\phi,\eta}$ are proper subsets of Ω_m. In what follows, when we speak of results about such sets as being true for $\eta \in {}^*\mathbb{N}\backslash\mathbb{N}$, this should be understood as restricted to appropriately small values of η. No such restrictions on the η-mapping are needed for sets of types 3 and 4.

We can now define A_η inductively:

- If $A = A_1 \cap A_2$, then $A_\eta = (A_1)_\eta \cap (A_2)_\eta$.
- If $A = A_1 \cup A_2$, then $A_\eta = (A_1)_\eta \cup (A_2)_\eta$.
- If $A = \bar{B}$, then $A_\eta = \bar{B}_\eta$.

We omit the (somewhat tedious) proof that the η-mapping is well-defined. The key step is the following:

LEMMA 2. If $B_1, \ldots, B_k \in \beta$ and $B_1 \cap \cdots \cap B_k = \emptyset$, then $(B_1)_\eta \cap \cdots \cap (B_k)_\eta = \emptyset$.

From this result, basic set theoretic arguments suffice to show that A_η is well-defined for every $A \in \mathcal{A}_{\mathcal{G}}$.

6.2. *Definition of* μ_G

We define the measure μ_G of Section 4.2 by the following formula:

$$\mu_G(A) = {}^\circ\left[\frac{\mu(A_\eta \cap G_\eta)}{\mu(G_\eta)}\right]$$

for $A \in \mathcal{A}_{\mathcal{G}}$, whenever this number is constant for $\eta \in {}^*\mathbb{N}\backslash\mathbb{N}$. We shall say that $\mu_G(A)$ is defined if this condition is satisfied. Since μ and the η-mapping are well-defined, so is the function μ_G. Note that G is just a 'type 1' set, specifically A_ψ where $\psi(n) = 10^{n-1}$, so

$$G_\eta = \{(i, n)/i \leq \psi(n \wedge \eta)\}.$$

The following theorem states that $\mu_G(A)$ is defined for $A \in \mathcal{A}_{\mathcal{G}}$.

THEOREM 1. For every $A \in \mathcal{A}_{\mathcal{G}}$, there is a constant real number a such that for every $\eta \in {}^*\mathbb{N}\backslash\mathbb{N}$, $^\circ[\mu(A_\eta \cap G_\eta)/\mu(G_\eta)] = a$. That is, $\mu_G(A)$ is defined and equal to a.

The theorem is proved first for the four basic types of sets A in $\mathcal{A}_{\mathcal{G}}$, i.e., for sets A in β.

1. Suppose $A = A_\phi$, where $\phi'(n)/10^{n-1} \sim a$. Then $\phi'(n) = a10^{n-1} + \alpha(n)$, where $\alpha(n)/10^{n-1} \sim 0$. Since $A_\eta \cap G_\eta = \{(i, n)/i \leq \phi'(n \wedge \eta)\}$, we have (by the definition of μ in (15))

$$(18)\ \mu(A_\eta \cap G_\eta) = \sum_{n=1}^{\eta} \frac{\phi'(n)}{m}\left(\frac{35}{36}\right)^{n-1}\frac{1}{36} + \frac{\phi'(\eta)}{m}\left(\frac{35}{36}\right)^{\eta}$$

$$= \sum_{n=1}^{\eta} \frac{a \cdot 10^{n-1}}{m}\left(\frac{35}{36}\right)^{n-1}\frac{1}{36} + \frac{a \cdot 10^{\eta-1}}{m}\left(\frac{35}{36}\right)^{\eta}$$

$$+ \sum_{n=1}^{\eta} \frac{\alpha(n)}{m}\left(\frac{35}{36}\right)^{n-1}\frac{1}{36} + \frac{\alpha(\eta)}{m}\left(\frac{35}{36}\right)^{\eta}$$

$$= a\mu(G_\eta) + \sum_{n=1}^{\eta} \frac{\alpha(n)}{m}\left(\frac{35}{36}\right)^{n-1}\frac{1}{36} + \frac{\alpha(\eta)}{m}\left(\frac{35}{36}\right)^{\eta}.$$

Since $\alpha(n)/10^{n-1} \sim 0$, given any $\epsilon > 0$ we can find a finite M such that $\alpha(n) \leq \epsilon \cdot 10^{n-1}$ for $n \geq M$. It follows that the second and third terms in (18) sum to less than

$$\sum_{n=1}^{M} \frac{\alpha(n)}{m}\left(\frac{35}{36}\right)^{n-1}\frac{1}{36} + \epsilon \cdot \mu(G_\eta).$$

Hence, when we divide (18) by $\mu(G_\eta)$ and take the standard part, the result is just a. So $\mu_G(A)$ is defined and equals a.

2. If $A = \bar{A}_\phi$, where $\phi'(n)/10^{n-1} \sim a$, then it follows that $\mu_G(A)$ is defined and equals $1 - a$, since $(A_{\phi,\eta} \cap G_\eta) \cup (\bar{A}_{\phi,\eta} \cap G_\eta) = G_\eta$.
3. Suppose $A = B_S$, where $S \subseteq \mathbb{N}$ is finite or co-finite.

(a) If S is finite, then ${}^*S = S$, so that $A_\eta = B_{S,\eta} = B_S$. Now

$$B_S \cap G_\eta = \{(i, n)/i \leq 10^{n-1}, n \in S\}.$$

Suppose K is an upper bound for S, for some finite K. Then

$$\begin{aligned}\mu(A_\eta \cap G_\eta) &= \sum_{n \in S} \frac{10^{n-1}}{m} \left(\frac{35}{36}\right)^{n-1} \frac{1}{36} \\ &\leq \sum_{n=1}^{K} \frac{10^{n-1}}{m} \left(\frac{35}{36}\right)^{n-1} \frac{1}{36} \\ &= \frac{1}{36m} \left[\frac{(350/36)^K - 1}{(350/36) - 1}\right],\end{aligned}$$

and it is clear that ${}^\circ[\mu(A_\eta \cap G_\eta)/\mu(G_\eta)] = 0$ for every $\eta \in {}^*\mathbb{N}\backslash\mathbb{N}$. So $\mu_G(A) = 0$ in this case.

(b) If S is co-finite, then $S = \bar{T}$ for some finite T. Then

$$A_\eta = B_{S,\eta} = \{(i, n)/n \leq \eta, n \notin T\}.$$

So

$$\begin{aligned}A_\eta \cap G_\eta &= \{(i, n)/i \leq 10^{n-1}, n \leq \eta, n \notin T\} \\ &= \{(i, n)/i \leq 10^{n-1}, n \leq \eta\}\backslash\{(i, n)/i \leq 10^{n-1}, n \in T\} \\ &= (F_\eta \cap G_\eta)\backslash(B_{T,\eta} \cap G_\eta),\end{aligned}$$

where F_η is the 'unsafe' set F of Section 3.2. There, we showed that ${}^\circ[\mu_\eta(F_\eta \cap G_\eta)/\mu_\eta(G_\eta)] = 5/162$; the result also holds if we substitute μ for μ_η. Since T is finite, ${}^\circ[\mu(B_{T,\eta} \cap G_\eta)/\mu(G_\eta)] = 0$, as just shown. Hence, for this case, $\mu_G(A) = 5/162$.

4. Finally, suppose $A = \bar{B}_S$, where $S \subseteq \mathbb{N}$ is either finite or co-finite. Then it follows from the above results that $\mu_G(A) = 1$ if S is finite and 157/162 if S is co-finite.

The next step is to prove that Theorem 1 holds if A is a finite intersection of sets in β. Recall (from Lemma 1) that the intersection of any two sets of the same type is another set of the same type; it follows that we may assume that A is a finite intersection of at most four sets, one of each type.

Furthermore, if S is finite and B_S is in the intersection, then case 3 above shows that $\mu_G(A) = 0$. If S is a co-finite set, then $S \cap \bar{T}$ is co-finite if T is finite, and finite if T is co-finite; this shows that we may assume that one of the following cases holds: (i) the intersection contains no sets of type 3 or 4; (ii) the intersection contains exactly one type 3 set B_S where S is co-finite, and no type 4 set; (iii) the intersection contains exactly one type 4 set and no type 3 set. It is straightforward that if we can demonstrate that $\mu_G(A)$ is defined for the first two cases, then it is defined for the third as well. Thus it remains only to prove that Theorem 1 holds for cases (i) and (ii):

Case (i). Suppose $A = A_{\phi_1} \cap \bar{A}_{\phi_2}$, where both ϕ_1 and ϕ_2 are $\bigcirc\psi$. Then $X_\eta = A_{\phi_1,\eta} \cap G_\eta$ can be expressed as the disjoint union of $Y_\eta = A_\eta \cap G_\eta$ and $Z_\eta = A_{(\phi_1 \wedge \phi_2),\eta} \cap G_\eta$. So $\mu(X_\eta) = \mu(Y_\eta) + \mu(Z_\eta)$. Now suppose $\phi_1(n)/10^{n-1} \sim a_1$ and $\phi_2(n)/10^{n-1} \sim a_2$. It follows that $(\phi_1(n) \wedge \phi_2(n))/10^{n-1} \sim (a_1 \wedge a_2)$. So dividing the equation just proven by $\mu(G_\eta)$, taking standard parts, and applying the first part of the proof, we conclude that $\mu_G(A) = a_1 - (a_1 \wedge a_2)$.

Case (ii). By employing the method of Case (i), it suffices to show that $\mu_G(A)$ is defined if $A = A_\phi \cap B_S$ where S is co-finite. Let $S = \bar{T}$, where T is a finite subset of $\mathbb{N}$. Then we have

$$\begin{aligned} A_\eta \cap G_\eta &= A_{\phi,\eta} \cap B_{S,\eta} \cap G_\eta \\ &= \{(i,n)/i \le \phi'(n \wedge \eta), n \le \eta, n \notin T\}. \end{aligned}$$

So

$$\begin{aligned} \mu(A_\eta \cap G_\eta) &= \sum_{\substack{n=1 \\ n \notin T}}^{\eta} \frac{\phi'(n)}{m} \left(\frac{35}{36}\right)^{n-1} \frac{1}{36} \\ &= \sum_{n=1}^{\eta} \frac{\phi'(n)}{m} \left(\frac{35}{36}\right)^{n-1} \frac{1}{36} - \sum_{n \in T} \frac{\phi'(n)}{m} \left(\frac{35}{36}\right)^{n-1} \frac{1}{36} \end{aligned}$$

The subtracted portion, when divided by $\mu(G_n)$, clearly has standard part 0, by part 3(a) of the proof for sets in β. And assuming that $\phi'(n) = a \cdot 10^{n-1} + \alpha(n)$ where $\alpha(n)/10^{n-1} \sim 0$, the first part of the expression, when divided by $\mu(G_\eta)$, has standard part $\frac{5}{162}a$. So $\mu_G(A)$ is well-defined and equals $\frac{5}{162}a$. This shows that A_ϕ and B_S are independent with respect to μ_G. This seems correct, since the set A_ϕ of draft positions for George and the set B_S of possible game lengths can be specified independently.

We have shown that $\mu_G(A)$ is defined if A is one of the sets in the basis β or a finite intersection of sets in β. Theorem 1 now follows from two easily verified facts:

1. If $A = \cup_{i=1}^{k} A_i$ is a disjoint union, and $\mu_G(A_i)$ is defined for $i = 1, \ldots, k$, then $\mu_G(A)$ is defined and equals $\sum_{i=1}^{k} \mu_G(A_i)$.
2. Any set in $\mathcal{A}_{\mathcal{G}}$ may be written as a disjoint union of finite intersections of sets in β.

6.3. *Construction of $\mathcal{A}_{\mathcal{G}}$ and ν_G*

The algebra for the original game of Section 4.3, which we again call $\mathcal{A}_{\mathcal{G}}$, has as its basis β the four types of set:

1. $A_\phi = \{(i,n)/i \leq \phi(n) \text{ and } i \in \mathbb{N}\}$, where ϕ' is $\bigcirc\psi$.
2. $\bar{A}_\phi = \{(i,n)/i > \phi(n)\}$, where ϕ' is $\bigcirc\psi$.
3. $B_S = \{(i,n)/n \in S\}$, where $S \subseteq \mathbb{N}$ is finite or co-finite.
4. $\bar{B}_S = \{(i,n)/n \notin S\}$, where $S \subseteq \mathbb{N}$ is finite or co-finite.

The full algebra $\mathcal{A}_{\mathcal{G}}$ consists of all finite unions of finite intersections of sets in β. The only change from Section 4.2 is that the draft number i is always restricted to $\mathbb{N}$, which makes the type 2, 3 and 4 sets different.

We would like to map each of these sets under σ to the corresponding set in Section 4.2. For instance, $\sigma(A_\phi)$ should just be A_ϕ again; and $\sigma(B_S)$ should be the same as B_S, except that i is extended to range over $\bar{m}$ instead of just $\mathbb{N}$. As we have seen, the mapping σ is (so far) not well-defined. However, because Z is a set of measure 0, we can simply tack it on in defining the σ-image. For instance, $\sigma(A_\phi) = A_\phi \cup Z$. Then σ is well-defined on β and can be extended to all of $\mathcal{A}_{\mathcal{G}}$ by the inductive definition. The following result holds:

LEMMA 3. If $B_1, \ldots B_k \in \beta$ and $B_1 \cap \cdots \cap B_k = \emptyset$, then $\mu_G[\sigma(B_1) \cap \cdots \cap \sigma(B_k)] = 0$.

We may thus define, for A in the algebra $\mathcal{A}_{\mathcal{G}}$,

$$(19) \qquad \nu_G(A) = \mu_G(\sigma(A)).$$

ACKNOWLEDGEMENTS

We appreciate the help of Richard Johns, who first thought of the truncated finite game (of Section 3.1), and the assistance of Ed Perkins for suggestions on how to set up the outcome space in Section 4.1. We also thank all those who have huddled with us around tables in dark corners of hotel bars, discussing this paradox (especially Jamie Dreier, Alan Hájek and Luc Bovens). Finally, we are grateful to two anonymous referees for helpful suggestions.

NOTES

[1] Similar versions of the "shooting-room paradox" are described in Leslie (1996) and Eckhardt (1997). The original formulation of the paradox is due to Leslie, who formulated it to present one version of his doomsday argument: if the population is increasing at a geometric rate, then we should assign a high probability to our belonging to the final, doomed generation.

[2] Exception: one hundred per cent if double six is rolled on the first round – a detail omitted from the verse for rhythmical considerations.

[3] More precisely, we assume that George's subjective probabilities are in accord with the Principal Principle: $P(A/B \cdot ch(A/B) = p) = p$, where P is George's subjective probability, ch is chance, and P does not incorporate any 'inadmissible' information about the truth of A. See Lewis (1980).

[4] At any rate, members of the population are ignorant of the mechanism for assigning numbers, so that the draft assignments appear random to them.

[5] De Finetti (1975), p. 123.

[6] A non-measurable event is one to which it is impossible to assign a consistent measure. For example, using the standard Lebesgue measure on the [0, 1] interval, it follows from the axiom of choice that there exists a subset of [0, 1] that cannot consistently be assigned a measure. See Royden (1968), p. 63 ff.

[7] Or perhaps it is fixed arbitrarily. That is, within measure theory one can prove the existence of a function having all of the characteristics of conditional probability and assigning a value to $P(A/B)$, but there will be many such functions and they need not agree on the value of $P(A/B)$ when $P(B)$ is not of positive finite measure. See Billingsley (1996), p. 427.

[8] It should now be clear that this paper has an underlying political objective: to 'empower' the measurably challenged sets by letting them claim their rightful place in the world of conditionalization.

[9] Note that this assumption excludes certain versions of the shooting-room game. For example, the assumption of independence will fail if the executioner knows in advance when the game will end (perhaps he has pre-rolled the dice, and simply reveals the result of each roll as the participants enter the room) and rigs the draft numbers so that George is certain to participate. (Or rather, the assumption will fail if George and Tracy *believe* that this is so.) This version of the game is particularly relevant in connection with the Doomsday Argument; for further discussion of that argument and the independence assumption, see our paper (1999).

[10] Actually, there is one way. If Tracy sees a complete list of all the participants, then she knows not only that the game was finite, but also the length of the game. So she knows that L_n is true, for some n. In this case, her probability for George to have died is $P(R_n/G \cdot L_n)$, which is given by Equation (2) as $(r_n - r_{n-i})/r_n$. This number could equal 0.9 for finitely

many values of n, even though the ratios converge to 0 as $n \to \infty$. In any case, we have assumed that Tracy is ignorant of the actual length, and knows only that the game was finite.

[11] See Hurd and Loeb (1985), II.6 for a discussion of the internal/external distinction.

[12] See Hurd and Loeb (1985), I.6.

[13] As before, $\bar{m} = \{1, \ldots, m\}$.

[14] For G, this follows from the fact that G is a subset of $\cup_{p=1}^{\infty}\{(i, n)/i = p\}$. Each of these sets has μ-measure $1/m$, and hence $\hat{\mu}$-measure 0. We can show that $(\hat{\mu})(F) = 1$ by writing F as the countable union $\cup_{j=1}^{\infty}\{(i, n)/n = j\}$, since the jth set in this union has measure $(35/36)^{j-1} \cdot 1/36$.

[15] Although the measure is not countably additive, this is not to be expected. For F is the countable union of the sets L_n, and we want $\mu_G(F) = \frac{5}{162}$, but $\mu_G(L_n) = 0$ for each n.

[16] By ${}^*A$, we mean the *-transform of A.

[17] Non-standard measures can thus be used to solve the lottery problem raised by De Finetti. We first learned how to define a non-standard measure on $\mathbb{N}$ from Brian Skyrms via Alan Hájek.

[18] Formally, the first G is a subset of Ω and the second a subset of Ω_m, but in fact the two sets of points are identical.

[19] We depart here from Leslie (1996), who maintains that the difference depends on whether or not the dice tosses constitute a deterministic process. Leslie reasons that in the "fully deterministic" case, "you must expect disaster. Disaster is what will come to over 90 per cent of those who will ever have been in your situation" (p. 255). We believe that his argument amounts to this: if George believes that the set-up is deterministic, then he believes that the game is determined to end, and hence that it will end. As we have shown, Leslie is right in one sense: if George believes that the game will end, then his degree of belief that he will die given that he participates is indeed 90 per cent. However, George's inference from 'the set-up is deterministic' to 'the game is determined to end' would be fallacious. Determinism by itself tells George nothing about which of the a priori possible outcomes (among them the infinite games) will occur.

[20] Indeed, of worrying (0.9)/(1/36) $\approx$ 32 times too much.

[21] ϕ and ψ are extended to the hyperfinite integers via the *-transform.

[22] By ${}^*S$, we mean the *-transform of S.

REFERENCES

Bartha, P. and C. Hitchcock: 1999, 'No One Knows the Date or the Hour: An Unorthodox Application of the Rev. Bayes' Theorem', *Philosophy of Science* **66** (Proceedings), S339–S353.

Billingsley, P.: 1996, *Probability and Measure*, J. Wiley & Sons, New York, 3rd edition.

de Finetti, B.: 1975, *Theory of Probability*, vols. 1 and 2, Wiley, New York. Trans. Antonio Machí and Adrian Smith.

Eckhardt, W.: 1997, 'A Shooting-Room View of Doomsday', *Journal of Philosophy* **XCIII**(5), 244–259.

Hájek, A.: 1999, 'What Conditional Probability Could Not Be', forthcoming in *Synthese*.

Hurd, A. E. and P. A. Loeb: 1985, *An Introduction to Nonstandard Real Analysis*, Academic Press, Orlando.

Leslie, J.:1996, *The End of the World*, Routledge, New York.

Lewis, D.: 1980, 'A Subjectivist's Guide to Objective Chance', in Richard C. Jeffrey (ed.), *Studies in Inductive Logic and Probability*, University of California Press, Berkeley, pp. 263–293.

Loeb, P. A.: 1975, 'Conversion from Nonstandard to Standard Measure Spaces and Applications in Probability Theory', *Transactions of the American Mathematical Society* **211**, 113–122.

Royden, H. L.: 1968, *Real Analysis*, Macmillan, New York, 2nd edition.

Paul Bartha:
University of British Columbia
Department of Philosophy
1866 Main Mall, E-370
Vancouver, B.C.
V6T 1Z1
Canada
E-mail: bartha@interchange.ubc.ca

Christopher Hitchcock:
California Institute of Technology
Division of Humanities and Social Sciences
Mail-Code 101-40
Pasadena CA 91125
U.S.A.
E-mail: cricky@its.caltech.edu