Generating ambiguity in the laboratory[1]

Jack Douglas Stecher[2] John Dickhaut[3]

August 2007

[1]This paper grew out of our correspondence with John Hey and Mark Machina. We thank them, along with Jonas Andersson, Monica Capra, Robert Frank, Nikolaos Georgantzis, Tom Issaevitch, Jack Kareken, Terje Lensberg, Jostein Lillestøl, Charles Noussair, Charles Plott,Leif Sandal, Erik Sørensen, Bertil Tungodden, and participants at the Eighth Conference of the Society for the Advancement of Economic Theory and at the 38th meeting of the European Mathematical Psychology Group. We also thank Cris Calude and Mike Dinneen for providing us with the quantum bits. All errors are our own.

[2]Corresponding author. Norwegian School of Economics and Business Administration, Helleveien 30, N-5045 Bergen, Norway. Tel: +47 55 95 94 27, Fax: +47 55 95 93 20. E-mail: jack.stecher@nhh.no.

[3]Carlson School of Management, University of Minnesota, 321–19 Avenue South, Minneapolis, MN 55455. Tel: +1 612 624 9891, Fax: +1 612 626 1355. E-mail: jdickhaut@umn.edu.

Abstract

This paper presents a method for drawing random numbers from a distribution that is, in an appropriate sense, unknowable to the subjects and to the experimenter. We show that the distribution's support can be finite, compact, or unbounded; the only requirement is nondegeneracy (i.e., the support must have at least two elements). We argue that these ambiguously-distributed random numbers can arise in natural settings, and discuss the implications for interpreting Ellsberg's experiments. We close with practical suggestions for experimenters using our technique.

Keywords: Ambiguity; laboratory experiments; methodology; Knightian uncertainty; Ellsberg's paradox

JEL Classifications: C90; C91; C92; D80; D81

1 Introduction

This paper provides a practical method for generating random numbers from a distribution that is, in an appropriate sense, unknowable. Specifically, if F is the cumulative distribution function of such a random variable and x is any value in the support of F, then it is impossible for the experimenter—and, a fortiori, for the subjects—to estimate $F(x)$ consistently, irrespective of how many draws one observes. We refer to such random numbers as *ambiguously* distributed.[1]

The technique we propose is fully general. As we demonstrate below, it is possible to generate ambiguously-distributed random numbers with a support of as few as two elements, or with infinite (bounded or unbounded) support. If an experiment requires generating lotteries where some choices have known probabilities and others are unknown, one can easily do so.

Our methodology differs from the traditional approach to ambiguity, introduced in Ellsberg. Conventionally, the experimenter withholds data from the subjects, but privately knows the likelihood of each outcome. In our setting, subjects can have access to the same information that the experimenter has. In fact, the experimenter may benefit from allowing subjects to observe different sequences of draws that use our technique. Thus while ambiguity might be dismissed under the conventional approach as a lack of sophistication or a cognitive bias, this case is much more difficult to make in our setting.[2] We take up this issue below.

Our motivation is primarily to present a technique for use in economics experiments. Nevertheless, we demonstrate how ambiguously-distributed random numbers can arise outside the laboratory, using purely physical randomization devices. The physical sources only have two possible outcomes, and thus can be seen as coin flips. Our manipulations of the outcomes only involve basic arithmetic and only use rational numbers. Thus the ability to generate ambiguous random numbers does not depend on pseudorandomness or on idealized mathematical abstractions.

The structure of this paper is as follows: Section 2 shows the details of our method. Section 3 shows the results from using our technique, both with physical randomization devices and with simulations. Section 4 discusses alternative approaches, and provides interpretations of our results and recommendations for researchers intending to use our technique. Section 5 concludes.

2 Description of the Technique

The main idea is to work with an autoregressive (of order 2) Cauchy process. Our intention is to begin with Cauchy random variables, which are known to present difficulties for inference, and then destroy the remaining properties that a statistician could find useful. In some sense, the distribution from which we sample is perfectly known, as we are able to characterize it. We view the distribution as ambiguous, however, because we construct it in such a way that every order statistic is unknowable.

The Cauchy distribution is a convenient point of departure. It is given by the cumulative distribution function and density

$$F(x) = \frac{1}{\pi}\arctan\left(\frac{x-x_0}{\gamma}\right) + \frac{1}{2}; \qquad f(x) = \frac{1}{\pi\gamma\left[1+\left(\frac{x-x_0}{\gamma}\right)^2\right]}$$

respectively. The parameter x_0 is called the *location* of the distribution, and the parameter γ is called the *scale*. These fully characterize the distribution. Among the properties of the Cauchy, it is known that, for any $\widehat{x} \in \mathbb{R}$,

$$E[x|x \leq \widehat{x}] = -\infty \text{ and } E[x|x > \widehat{x}] = +\infty.$$

Since one can make $\infty - \infty$ equal anything, it is easy to see that there is no mean of a Cauchy-distributed random variable; this is true irrespective of the location and scale. It is also known that all higher moments about any number are undefined; for example, the second moment (hence any replacement for variance) of a Cauchy-distributed random variable is infinite. As the hypotheses of finite means and variance underlie laws of large numbers and central limit theorems, it is unsurprising that neither holds for a Cauchy random variable. This alone makes statistical inference difficult in the Cauchy case.

Generating approximately Cauchy-distributed random variables is entirely straightforward. Laha (1959) provides a large number of bivariate distributions $G(X, Y)$ for which the ratio X/Y has a Cauchy distribution. In particular, for any two independent normally-distributed random variables X, Y with $E[X] = E[Y] = 0$, X/Y is Cauchy-distributed with location 0. Accordingly, if one can generate independent normally-distributed random variables and divide, then one can obtain Cauchy-distributed random variables.

Despite the failure of laws of large numbers and central limit theorems, and despite the lack of any well-defined moments , the Cauchy distribution will not serve our purpose without further modification: Barnett (1966) shows that the sample order statistics are good estimators of the Cauchy distribution's theoretical quantiles. So while the usual descriptive statistics may be beyond reach, the cumulative distribution function (and hence whatever probabilities one wants) can be estimated. In fact, the sample median is a good estimator of the location x_0. To get a distribution where inference is futile, we need to do more.

The first change we make is to take advantage of the fact that a Cauchy-distributed random variable has no mean. We therefore define a first-order autoregressive process X_t by

$$X_0 \sim C[0,1] \qquad X_{t+1} \sim C[x_t, 1],$$

where we adopt the convention of using lower-case letters to indicate realizations and capital letters to indicate random variables. In words, we draw our first random variable from a Cauchy distribution with a location of 0 and a scale of 1 (called a *standard* Cauchy). For each subsequent random number, we draw from a Cauchy distribution with a location equal to the current realized value and with a scale still equal to one. Since the location of the next draw is the realization of the current one, it easily follows that the location of any future draw has no mean (as it would be the mean of a Cauchy-distributed random variable, which does not exist).

The above procedure still does not suffice to attain our goal. This may seem surprising, as all but the first draw come from Cauchy distributions with ex ante unknowable locations. However, drawing from a Cauchy distribution with location x_0 and scale 1 is equivalent to drawing from a standard Cauchy and then adding x_0; one simply has a shift of the origin. Pitman and Williams (1967) show that many functions (including sums) of Cauchy-distributed random variables are surprisingly well-behaved. In the particular case of interest, the problem that arises is essentially that the sum of n standard Cauchy random variables is itself Cauchy, and the location of the sum is still 0. Thus the point Barnett raises remains: the sample order statistics provide insights on the shape of the underlying distribution.

However, the above thought experiment puts us on the right track. We have so far used the Cauchy distribution to modify one of two parameters, the location. With a second-order autoregressive

process, we can do the same thing to the scale parameter γ. This enables us to generate a process that does not satisfy the requirements of Pitman and Williams; we can then show that the resulting process indeed does not have well-behaved order statistics. The only caveats are that, first, we must make certain that the scale parameter is always positive, and second, we may wish to prevent the scale from becoming too large too rapidly. To see this latter point, note that the reason that a Cauchy-distributed random variable has no mean is that extreme values occur with relatively high probability. If the scale of a Cauchy distribution is chosen as a large realization from a previous Cauchy draw, the likelihood of obtaining an extreme value will increase, causing the autoregressive Cauchy process to generate values that will quickly cause overflows in many software applications.

Accordingly, the main building block of our procedure works as follows:

- Draw $Z_0 \sim C[0, 1]$.
- Draw $Z_1 \sim C[z_0, 1]$.
- Let $\phi, \psi \in (0, 1)$ with $\phi, \psi \approx 0$. For $n \geq 2$, draw $Z_n \sim C[z_{n-1}, \phi \cdot |z_{n-2}| + \psi]$.
- Once a sample of the desired length is produced, resample the resulting data to destroy the autoregressive structure.

The parameters ϕ and ψ assure that the scale parameter is non-zero. Picking both close to zero slows the tendency of the resulting data to diverge toward $\pm\infty$.

The last step in the above procedure is necessary to prevent the data from being predictable. After the draws $\langle z_{n-2}, z_{n-1} \rangle$ are realized, an experimenter would know the distribution of the next draw, as it is simply a Cauchy with the realized location and scale. To make this information unavailable, it suffices to shuffle the data, with replacement. For a sample of length m, one need only draw a random sequence of m integers from $\{1, \ldots, m\}$, say from a discrete uniform distribution. One can then use the resulting sequence to choose observations from the first m terms of the AR(2) Cauchy process z. The resulting sample is distributed according to the empirical histogram of z but lacks the autoregressive structure, thereby destroying the few clues for inference that had remained in z.

The procedure just described generates a sequence of random draws where the scale shifts about an

unstable location. In the next section, we show that this indeed produces a data set with meaningless sample order statistics. That is, it is impossible to estimate consistently the distribution $F(x)$ for any $x \in \mathbb{R}$. This suffices for the case where an experimenter desires ambiguously-distributed variables with unbounded support.

To generate ambiguously-distributed random numbers with finite or compact support, one need only make a small modification. This result surprising, since a distribution with finite support or finite range cannot have infinite variance, and the infinite variance is a driving force in generating the instability of the AR(2) Cauchy process just described. Yet the structure of the Cauchy process does more than is at first apparent. Because the scale shifts, and because we dampen the scale parameter, our technique frequently produces sequences of draws that are close to each other. On the other hand, because the Cauchy distribution with any scale frequently generates extreme outliers, our technique will also create a large number of realizations that are arbitrarily far apart. This feature is essential to obtaining the desired properties in the harder cases of compact and finite support.

The procedure for generating ambiguously-distributed numbers with finite support—say, with k elements—works as follows:

- Generate a sequence $\{z_n\}$ by the AR(2) Cauchy procedure described above.
- For each n, let $y_n = \text{Floor}(z_n) \bmod k$.
- Once a sample of the desired length is produced, resample the resulting data to destroy the autoregressive structure.

As an example, imagine that $z_n = -23.7894$, and that $k = 10$. The floor of z_n is the largest integer that is no larger than z_n, so in this case, $\text{Floor}(z_n) = -24$. One then takes the remainder of dividing -24 by 10. Since -30 is the largest multiple of 10 that is no larger than -24, we calculate the remainder as $-24 - (-30) = 6$. If instead z_n has been positive 23.7894, we would have obtained $23 \bmod 10 = 3$. Thus we always end up with results in $\{0, \ldots, k-1\}$. With a value of ϕ close enough to 0, this procedure generates many observations where the current realization is within one integer of the previous one; however, because of the high number of extreme values in

a Cauchy distribution, this property does not persist. Shuffling of the resulting set of realizations then eliminates the ability of subjects or of the experimenter to use the underlying autoregressive structure as a basis for inference.

One can use the same general principle to generate ambiguous draws with compact support. Let $[a, b] = \text{supp}(F)$. Without loss of generality, we restrict attention to the case where a and b are both integers; if this is not the case, one can always perform an appropriate rescaling. The procedure works as follows:

- Generate a sequence $\{z_n\}$ by the AR(2) Cauchy procedure described above.
- For each n, let $y_n = \text{Floor}(z_n)$, and let $y'_n = z_n - y_n$. This means that y'_n is the decimal portion of z_n and y_n is the greatest integer less than or equal to z_n.
- Let $y''_n = a + [y_n \bmod (b - a)] + y'$.
- Once a sample of the desired length is produced, resample the resulting data to destroy the autoregressive structure.

To be precise, the above only generates ambiguous random variables from the half-open interval $[a, b)$. To include the right endpoint, we need an additional step: each time the left endpoint is drawn, make an auxiliary ambiguous draw from $\{0, 1\}$, using the procedure we provide for drawing from a distribution with finite support. If a 0 is selected, keep $y''_n = a$, but if a 1 is selected, instead change the value of y''_n to b. Now the full compact interval $[a, b]$ can appear.

It should be noted that, in the cases of finite and compact support, we can say something about the cumulative distribution function F: we know that, if $a = \min(\text{supp F})$ and $b = \max(\text{supp F})$, then of course $(\forall x < a)\ F(x) = 0$ and $(\forall x \geq b)\ F(b) = 1$. However, this is all that we can say. One cannot assert any probabilistic statements other than those that follow from the definition of what is in principle possible.

3 Results

We follow two approaches to illustrate our technique. First, we show the results of simulations using random number generators. Doing so enables us to consider large numbers of random draws and to demonstrate that the sample order statistics of the ambiguously-distributed numbers are indeed divergent.

Second, we use physical randomization devices in order to build up ambiguously-distributed random numbers. This illustrates that our results are not an artifact of mathematical idealizations or software architecture, but rather a naturally occurring phenomenon. All of the steps we use in this second approach are ones that can occur naturally: we restrict ourselves to the simple arithmetic operations of addition, subtraction, multiplication, and division, and use Bernoulli random variables as our building blocks. The resulting data are ambiguously distributed. Thus, if one can face both Bernoulli random variables and arithmetic, one can end up with ambiguity.

3.1 Simulations

For the simulations, we drew 10,000 ambiguously-distributed random variables, both with unbounded support and with support in $\{0, \ldots, 9\}$. We then replicated the procedure 100 times. We chose $\phi = 0.001$ and $\psi = 0.0001$.

For each unbounded sequences of 10,000 draws, we computed several order statistics, and then compared how the sample order statistics behaved over the 100 replications. The range of values for each sample order statistic, along with the median and mean values, are shown in Table 1.[3]

It is clear from Table 1 that none of the sample order statistics is convergent, after $10,000 \cdot 100 =$ 1 million draws. For example, the minimum of 10,000 draws was below -3 million in one replication and was strictly positive (just over 9) in another. The maximum ranged from below -2 to above 18 trillion. Thus there was at least one replication where all 10,000 draws were negative, and at least one replication where all 10,000 draws were positive. The most stable of the sample order statistics was the 10 %ile, where the highest value exceeded the lowest by more than 54,000. Finally,

Order Statistic	*Minimum*	*Median*	*Mean*	*Maximum*
Minimum	$-3.32017 \cdot 10^6$	-66.2505	-39718.8	9.20083
1 %ile	-562905	-51.9015	-11119.7	11.292
2.5 %ile	-145903	-40.9968	-4914.81	12.016
5 %ile	-96887	-29.9379	-3774.48	23.3314
10 %ile	-54679.7	-10.7992	-1172.47	36.0455
Lower Quartile	-33207.8	-1.45783	53.3294	47618.7
Median	-5348.3	0.0298731	2963.11	296289
Upper Quartile	-1196.73	1.70167	$1.66938 \cdot 10^7$	$1.66926 \cdot 10^9$
90 %ile	-371.732	7.91172	$7.09551 \cdot 10^7$	$7.09474 \cdot 10^9$
95 %ile	-342.251	19.2627	$1.06024 \cdot 10^8$	$1.06006 \cdot 10^{10}$
97.5 %ile	-27.9362	30.8853	$1.20062 \cdot 10^8$	$1.20025 \cdot 10^{10}$
99 %ile	-24.6536	39.7169	$1.33105 \cdot 10^8$	$1.33064 \cdot 10^{10}$
Maximum	-2.16994	46.6691	$1.88546 \cdot 10^8$	$1.88503 \cdot 10^{10}$

Table 1: Sample order statistics, 100 replications of 10,000 draws

although Barnett had found that the median was a good estimator of the location of a Cauchy-distributed random variable, the same cannot be said here: the sample median ranged from under -5000 to nearly 300,000.

While the sample order statistics of the AR(2) Cauchy process are plainly divergent, the autoregressive structure remains. Figures 1–3 provide graphical representations of three additional replications (i.e., not included in generating the data in Table 1).

[Figure 1 about here.]

In Figure 1, all 10,000 realized values are in the interval $(-40, 30)$. For the first 5,000 draws, all the realizations are between -5 and 5. The realizations above 5 all occur in the next 1,500 draws, after which all the realizations below -5 occur. For the last 2,000 draws, every realization is again between -5 and 5.

[Figure 2 about here.]

Contrast this picture with Figure 2, where the range of realized values is much more narrow: every draw is in the interval $(-2, 10)$. The most extreme realizations occur in the early draws, with all those above 3 or below -1arriving within the first 2,000 draws. From draw 4,000 onward, every realization is close to zero.

[Figure 3 about here.]

Figure 3 shows a replication with a much wider range of values, and where the extreme values occur in the latter portion of the time series. The realized values range from just below $-4,000$ to a high of around 200. The first 5,000 draws contain all of the positive realizations. The most extreme observations, at $-2,000$ or below, occur at two points in the time series, either just after draw 6,000 or between draws 9,000 and 10,000. The bulk of the last 5,000 draws fluctuate around $-1,000$, with dramatic variation compared with that of the first 5,000 draws.

As expected, each replication is visibly from a seemingly different distribution, but one can also easily see the autoregressive structure in the data. We point this out to emphasize the importance of shuffling the data after generating a sample of desired length, as suggested in Section 3.

To illustrate the case of restricted support, we converted the simulated draws just described to the finite range $\{0, \ldots, 9\}$, using the procedure described in Section 2. In Table 2, we show the results of the first ten replications. Again, each replication consisted of 10,000 draws.

Replication	0	1	2	3	4	5	6	7	8	9
I	28	499	795	783	1,520	5,593	512	36	0	234
II	1,016	837	483	573	244	46	768	2,477	1,467	2,089
III	1,521	639	251	575	902	827	1,042	2,323	974	946
IV	0	0	0	0	0	0	54	2,084	6,833	1,029
V	2,220	236	580	603	632	912	758	550	378	3,131
VI	666	712	751	842	1,725	1,641	1,191	707	845	920
VII	17	119	187	483	7,952	754	292	31	102	63
VIII	315	360	757	5,901	495	723	360	226	569	294
IX	2,261	1,034	378	90	77	221	92	152	79	5,616
X	1,795	1,658	861	298	106	302	303	32	1,920	2,725

Table 2: Sample histogram, finite support, 10 replications

Observe that in the first replication, the most commonly drawn digit was 6, making up 55.93% of the realizations, while 8 was the least common draw (and in fact did not occur at all in 10,000 tries). However, on the second replication, 14.67% of the draws (just over 1/7) were equal to 8, whereas only 0.46% were equal to 6. Note that these were not continuations of the same autoregressive process: for each replication, we restarted.

The results do not stabilize over the full set of 100 replications, and instead become even more dramatically unstable. For each digit in $\{0, \ldots, 9\}$, there was at least one replication in which it had zero realizations out of 10,000 draws, and there was at least one replication when it made up at least 28.5% of the draws. Eight of the ten digits had at least one replication where it made up more than half of the draws. The most extreme case (our 43^{rd} replication) had 88.92% of realized values equal to 0, 11.08% equal to 1, and no other values realized. The most symmetric case (our 66^{th} replication) had each digit occur at least 8.8% of the time and at most 11.47%. Even this case, however, is inconsistent with the data coming from a discrete uniform distribution (at any conventional significance level).

3.2 Physical randomization

We now show that ambiguously-distributed random numbers can arise in physical settings, without reliance on idealized mathematical properties of a Cauchy distribution or on the particular pseudorandom number generator found in a given software application.

We began by generating a large number of quantum bits (qbits). The procedure works as follows: we fired a photon at a semi-transparent mirror. On the opposite side of the mirror, a device attempted to detect the photon. If the device succeeded, a qbit was assigned a value of 1; otherwise, it was assigned a value of 0.

It is not universally accepted that a qbit generated in this way has a stationary or even well-defined probability of success.[4] In that case, generating qbits alone suffices to produce ambiguously-distributed random numbers. Hence we only need to show that we can produce ambiguously-distributed random numbers with qbits that are Bernoulli-distributed with success probability p.

We generated a sequence of 30 qbits, summed the result, and subtracted 14.99. The sum of 30 independent Bernoulli draws from the same distribution is approximately normal; this is just the Lindberg-Levy central limit theorem. If p is close to 1/2, then the expectation of the sum of 30 Bernoulli draws is approximately 15, and in any case, the sum is necessarily a nonnegative integer. Subtracting 14.99 therefore produced a number that could not have been 0, but if p is not too far from 1/2, then the resulting sum would have a mean reasonably close to 0.

We then repeated the process with the next 30 qbits we obtain. This gave us two random variables that were approximately normally-distributed with an identical mean, which is empirically close to zero. By Laha's argument, the ratio of these two random variables is therefore approximately Cauchy-distributed. Hence with 60 qbits, we generated a random number which is either already ambiguously distributed or which has an approximate Cauchy distribution. Because we subtracted 14.99 from an integer, we never have to concern ourselves with the possibility of division by 0.

From here it was straightforward to generate a AR(2) approximately Cauchy process, as described in Section 2. The first 60 qbits provided a Cauchy-distributed random variable with some location x_0 and scale γ. Call this realization z_1. To generate the next term in the process, we repeated the procedure with the next 60 qbits, then added z_1 to the result; this gave a Cauchy-distributed random number with location $x_0 + z_1$ and scale γ. For the n^{th} collection of 60 qbits, we generated a new Cauchy-distributed random variable, multiplied by the absolute value of z_{n-2}, multiplied the result by 0.01, and added z_{n-1}. That is, we chose $\phi = 0.01$.

After generating a sequence of 30 variables by this AR(2) approximately Cauchy process, we mapped the result to the interval $\{1, \ldots, 10\}$ using the procedure in Section 2 (and adding one to the result, solely so that we could start counting from 1 rather than from 0). We then repeated the procedure two more times, starting each replication over rather than continuing the existing process.

[Figures 4–6 about here]

Figures 4–6 show the resulting empirical histograms. Observe that in the first replication, the most common realized value was 3 (making up 30% of the draws), which did not show up at all in either the second or third replication. Summing across all three empirical histograms, the number 2 was never drawn; 8 was only drawn once (in the second replication); 6 was drawn twice (once each in the first and second replications). At the opposite extreme, 24.4% of the draws (22/90) were equal to 4, and 35.6% (32/90) were equal to 9. Thus overall these two outcomes accounted for 60% of the realizations.

To test the most extreme case of finite support, we mapped the same three AR(2) approximately Cauchy sequences to $\{0, 1\}$, again using the technique in Section 2. The first replication had an empirical success frequency of 23.3%; the third was 30%. Both of these differed significantly from

a success probability of 1/2 (with 99.5% confidence for replication one and with 95% confidence for replication three). For replication two, the empirical success frequency was 53.3%, which cannot be distinguished from a success probability of 1/2 at any defensible significance level. In sum, it is quite clear that the success probability in the second replication is different from that in the first and third, even though we followed an identical procedure.

4 Discussion

4.1 Alternative approaches

We begin by discussing two alternative approaches to ours. The first, which Hayashi and Wada (2006) use, involves a separate group of subjects whose task is to write down a distribution of balls in an urn under various constraints. Hayashi and Wada then randomly select one of these proposed distributions; this procedure is common knowledge among all subjects. Thus, subjects in the latter stage of their experiment (an Ellsberg-inspired task) are made aware that the experimenter does not know in advance what the distribution of balls in the urn will be.

Hayashi and Wada acknowledge several shortcomings of their approach, despite its advantage of expositional clarity. They note (and show graphically) that the subjects who are to mix balls overwhelmingly choose uniform distributions. Among those subjects who do not choose uniformity, there is some tendency to choose extremes (i.e., all or nearly all balls of the same color). Thus while the goal is to generate ambiguity, the result is largely predictable.

A second approach, close in spirit to ours, is in Calude et al. (2002).[5] Their idea, based on the work of Chaitin (1998), is to take advantage of the impossibility of solving Turing's halting problem. Chaitin shows that the unsolvability of the halting problem implies that the probability of a Turing machine halting is also uncomputable. Using this result, Calude et al. design some simple programs, which are special cases in that they will either halt in some knowable number of steps or run forever. Running such a program is thus equivalent to drawing a random number from a binomial distribution, where the success probability is unknowable.

Several factors make the specific technique in Calude et al. impractical for economics experiments. Each random number one wishes to draw requires a special program, in particular one for which it is possible to determine whether the program will halt. (The Turing halting problem only says that this cannot be determined for arbitrary programs.) The number of such programs that one can design for a given universal Turing machine is limited, and thus experimenters would need to design new universal Turing machines each time the supply of these special programs gets exhausted. Moreover, in the case where this is not an issue, there are also difficulties. Suppose that, for a given machine, there is a large number of programs known to halt within n steps if they ever halt. One might use this information to estimate the halting probability of the machine. Calude et al. do precisely this, and are able to obtain an exact value for the first 64 bits of the halting probability. So while this probability is technically unknowable, it can be estimated with extremely high precision.

4.2 Comments on our approach

We make use of two opposing forces in the AR(2) Cauchy process. The properties of the Cauchy distribution make extreme draws occur with a relatively high frequency. However, the coefficients ϕ and ψ on the scale parameter γ, if sufficiently small, have a damping effect, making a large number of draws stay close together.

The purpose of the additive constant ψ is simply to avoid division by 0, so we concentrate here on the effects that ϕ has. As ϕ increases, the magnitudes of relatively frequent draws grows rapidly. To illustrate, we simulated 100 replications of 10,000 draws, this time with $\phi = 0.1$ instead of the value of 0.001 used in Section 3. We kept $\psi = 0.0001$ as before.

Because these simulations damped less, the behavior of the unbounded draws was far more dramatic than in the simulations of Section 3. Table 3 shows the range of the mean and median over the 100 replications, each to one decimal place for the purpose of illustration:

All the other order statistics are similarly erratic. It is clear that data of these absolute magnitudes can easily generate overflows in some systems, and may be impractical for many computational

Statistic	*Minimum*	*Median*	*Maximum*
Median	$-1.6 \cdot 10^{17}$	0	$1.1 \cdot 10^{19}$
Mean	$-2.3 \cdot 10^{122}$	$-2.5 \cdot 10^{47}$	$5.1 \cdot 10^{110}$

Table 3: Sample mean and median, $\phi = 0.1$

purposes. More importantly, however, is that data of these absolute magnitudes are highly unlikely to be stored as an integer within a software application. One may still take the floor of a number such as $5.1 \cdot 10^{110}$ and then take the remainder after integer division. However, the data generating process is fundamentally different from the one we propose, and the behavior of the data are likely to introduce idiosyncrasies of the particular software one is using.[6]

As a further test, we computed 100 more replications of 10,000 draws, this time increasing ϕ to 0.2. The range of the median increased further, from a low of $-4.1 \cdot 10^{25}$ to a high of $5.0 \cdot 10^{29}$. To illustrate further, we note that the minimum ranged from $-7.4 \cdot 10^{294}$ to $-4.2 \cdot 10^{179}$. Thus, setting ϕ to 0.2 seems to have been inviting values close to those that would cause overflows. In sum, setting the damping parameter to 0.1 or 0.2 does increase the magnitude of the fluctuations in the unbounded data. However, Section 3 demonstrated that severe fluctuations were already attainable with $\phi = 0.001$, without risking overflows or making the conversion to integers questionable.

Just as increasing the damping parameter ϕ causes greater instability in the underlying AR(2) Cauchy process, reducing ϕ makes the frequency of extreme values relatively low. To illustrate this, we drew another 100 replications of 10,000 ambiguously-distributed numbers, this time setting $\phi = 0.000001$. Since overflows are not a concern in this case, we analyze the setting of finite support, specifically of support in $\{0, \ldots, 9\}$.

Among the 100 replications, there were 82 where at least 9,999 out of 10,000 draws all ended up as the same digit (though the specific digit varied from replication to replication). There were another 6 replications where two digits accounted for all the realizations, and 8 more where three digits accounted for all the realizations. The remaining four replications included one where four digits accounted for all the draws, one where six digits accounted for all the draws, and two where all ten digits had some realizations. In one of these last two cases, each digit was realized at least 517 times and at most 1532 times; the other case was slightly less symmetric but similar.

The impact of the small damping parameter is thus to confine the draws to a narrow range (in all but two out of 100 replications). There is evidently a trade-off between a low damping parameter, which keeps the realized values from getting too extreme and causing overflows, and a larger one, which keeps the realized values from being concentrated on a narrow range of outcomes. Yet it is noteworthy that even in the case of an extremely low damping parameter, there was a dramatic difference across replications, reinforcing the idea that the underlying distribution of the data is ambiguous.

4.3 Remarks on Ellsberg and advice to experimenters using our technique

The examples in this paper have concentrated on the case of complete ignorance, but it is possible to use our technique in settings where some probabilistic information is available. As motivation, we discuss the case studied in Ellsberg (1961).

Subjects in Ellsberg's experiments were to choose between two pairs of lotteries over the same prize. In one task, the subjects were to choose between a lottery giving the prize with probability $1/3$ and another giving the prize with an ambiguous probability in $[0, 2/3]$. In the second task, the subjects were to choose between a lottery giving the prize with probability $2/3$ and another giving the prize with ambiguous probability in $[2/3, 1]$. The main force behind the "paradox" is that the subjects do not act as if they assign a unique probability to the ambiguous events. This point is addressed theoretically in the literature, for example in Gilboa and Schmeidler (1989) and Ghirardato et al. (2004). For our purposes, it is worth noting that nothing would seem paradoxical about not assigning a probability, given that there are situations where one cannot meaningfully do so. Indeed, Savage (1954) himself acknowledges that he does not intend his subjective expected utility theory to apply to such contexts.[7]

There are at least two ways that an experimenter can use our technique to generate the situation specified in Ellsberg's experiments. First, one can use our technique to draw an ambiguous random number p from the compact interval $[0, 2/3]$. Call the resulting number the probability of drawing a black ball from the urn, and keep the probability of drawing a red ball equal to $1/3$ (this was the event with a known probability). The remaining probability $2/3 - p$ is then the probability of

the third event (drawing a yellow ball from the urn). Alternatively, a second approach can note that Ellsberg specified an urn with 90 balls, of which 30 are red and 60 are black or yellow. The experimenter can therefore draw an ambiguously-distributed number from $\{0, \ldots, 60\}$ in order to determine the number of black balls in the urn, and then make the remainder yellow.

The procedure just sketched illustrates that one can combine ambiguously-distributed numbers with those drawn from a fixed distribution in order to create partially ambiguous draws. In the case of the Ellsberg experiment, the argument would be especially compelling if the experimenter could convince the subjects that the ambiguous distribution is indeed unknowable.

One way to persuade subjects that they face ambiguity is to present them with information similar to what we present in this paper. By allowing subjects to see the results of different replications of the same process and to draw ambiguously-distributed numbers themselves, one could make readily apparent that there is no basis for inference. Direct experience seems to us more likely to be convincing than an argument or assertion about a probability being unknowable. At a minimum, an experimenter may wish to include several empirical histograms showing different replications of the same process. In this way, the inconsistency across pictures of a consistently used procedure could clarify the message.

5 Conclusion

We provide a technique for drawing random numbers from a distribution that is truly ambiguous. If a researcher wishes to draw ambiguously-distributed numbers with finite support, our technique can guarantee that none of the possible outcomes can have a known probability, and that no law of large numbers or central limit theorem holds (thereby ruling out any estimation from empirical sampling). A trivial modification enables an experimenter to draw from partially ambiguous distributions, as in the experiments of Ellsberg. If instead the experimenter wishes to draw from a distribution with infinite support, our technique makes the cumulative distribution function unknowable for any point in the support (with only the trivial exceptions of the endpoints in the case where the support is bounded).

The ability to draw ambiguous random numbers is not an artifact of a mathematical idealization, nor can it be attributed to the peculiarities of computer pseudorandom number generators. We demonstrate this by using physical randomization devices with binary outcomes. If these are seen as Bernoulli random variables with a stationary success probability, then an experimenter can generate ambiguous random variables simply by adding, subtracting, multiplying, and dividing the realized values, along with addition of and multiplication by two appropriate constants (both of which are rational numbers in our demonstrations). If on the other hand one treats the physical randomization devices as having a nonstationary or indeterminate probability, then ambiguity is already present.

For experimenters, the ability to generate ambiguous random numbers creates the possibility of studying settings where expected utility maximization does not apply. Doing so requires persuading subjects that the numbers being drawn come from an unknowable distribution. Replicating sampling from a given distribution and showing the resulting histograms to the subjects might help in this regard, as might allowing subjects to sample directly from the ambiguous distributions. If the results in this paper are persuasive, then providing similar demonstrations seems to be a reasonable way to try to convince subjects that they cannot rely on a law of large numbers.

Some possible applications may include models with non-unique expected utilities; indeed, we do precisely this (with a different methodology) in Stecher et al. (2006). The non-unique priors of Gilboa and Schmeidler (1989) and the imprecise probabilities of Koopman (1940a,b, 1941) all fit into this setting, as do models of ambiguity aversion, such as Ghirardato et al. (2004), Maccheroni et al. (2006), or Bossaerts et al. (2007).

There are some technical issues of which an experimenter should be aware before using our technique. First, because our technique uses an autoregressive structure, the sampling procedure should be designed in such a way as to destroy this structure. Otherwise, subjects could draw inferences from the realized time series. To avoid this issue, we recommend a two-step procedure: first, draw a sample of the desired length, using the technique described in Section 3. Then scramble the data, using a bootstrap-like or shuffle-like method.[8]

Second, there are questions about setting the scaling parameter ϕ in our procedure. We recommend

setting ϕ close to zero, and indeed show that, with values of ϕ of 0.1, the realized draws are of such great magnitude that they are unlikely to be stored as integers. This is unproblematic when the experimenter is only concerned with drawing ambiguous random numbers with unbounded support. For any other cases, however, having ϕ this large will likely mean that the resulting data set will not have the properties we describe, but instead will reflect the techniques a particular software application uses for storing large numbers. Even in the unproblematic case of unbounded support, we saw that having ϕ near 0.2 resulted in realized values on the order of 10^{294}. At these magnitudes, a researcher would need to become concerned about overflows in computer memory.

The results we presented instead had $\phi = 0.001$ in our simulations and $\phi = 0.01$ in our experiments with physical randomization devices. Both produced the ambiguous data as desired. As a test, we set ϕ much closer to 0, choosing 0.000001. We found that, while the resulting data sets were ambiguous, in the sense that they varied wildly from replication to replication, over 80% of the replications with finite support had 99.99% of the realized values equal the same number. (The number, however, was different across replications.) The reason is that ϕ is used in scaling, and there was simply too subtle variation in the realized numbers for our technique to detect in most replications. There were, however, several cases where even this small value of ϕ produced an ambiguous distribution with more dispersion. In 2% of our replications, the resulting distribution had every value in the support realized a significant number of times. However, we view this frequency of achieving dispersion as too low to be acceptable. Accordingly, we recommend that the value assigned to ϕ be in the range of our main results, that is, between 0.001 and 0.01. These levels seem to attain a large amount of dispersion in the output without generating severely explosive behavior. As our intention is for our technique to find use in economics experiments, it is our hope that research using our technique may shed additional light on optimally setting this parameter.

6 Notes

Section 1

1. Our notion of ambiguity differs from that in Ellsberg (1961): the latter does not preclude subjects from forming a meaningful or even correct belief about the distribution of balls in an urn, whereas our technique does. In this sense, what we draw is closer to Knightian uncertainty (Knight 1921). The notion of *complete ignorance*, under which there is no probabilistic knowledge over the (fully known) set of possible states, is also clearly related. Complete ignorance appears at least as far back as Hurwicz (1951), and is discussed in Luce and Raiffa (1957), though they describe the concept's definition as "somewhat vague;" other work using complete ignorance includes Pazner and Schmeidler (1975), and more recently Bossert and Slinko (2006) and a related working paper by Baigent (2007). The technique we employ focuses on the case of complete ignorance but need not be so restrictive—we can easily adapt our technique to give some limited probabilistic knowledge (called *partial ambiguity* in Rustichini et al. (2005)). In the interest of having a consistent term to use in structuring the argument, we choose ambiguity, and use *true ambiguity* when we wish to emphasize the distinction between our situation and that of Ellsberg. Vierø (2006, 2007) uses the term *vagueness*, but we prefer ambiguity, as vagueness has other connotations in the philosophy literature. The term ambiguity is used to denote a primitive concept elsewhere in the literature, for example in Fishburn (1993) and Nehring (2000, 2007).

2. There is widespread evidence that subjects view ambiguity differently from how they view settings where they face a known risk. An example from the neuroeconomics literature is Rustichini et al. (2005), who find (among other things) that subjects have shorter response times when deciding between ambiguous and certain lotteries, compared with response times when risk is involved. Rustichini et al. also note numerous physiological differences between how subjects react to ambiguity and how they react to risk. Gilboa and Schmeidler (1989, 1993) and Schmeidler (1986, 1989) have proposed models of non-additive probabilities and of non-unique priors; in these settings, Ellsberg's subjects act in ways that are rational in the appropriately modified sense. Bewley (2002), Gärdenfors and Sahlin (1982, 1983), Koopman (1940a,b, 1941), Nehring (2007), and Stecher (2007a,b) are similar in spirit.

Section 3

3. These simulations were produced in *Mathematica*. Similar results were obtained using *R*, so we are confident that they are not due to the particular pseudorandom number generator used.

4. Calude (2004, p. 10) and Calude and Svozil (2006, p. 6) note that quantum random number generators do not have a success probability of 1/2, even when the output is passed through an unbiasing algorithm. Others argue that a probabilistic interpretation of quantum phenomena must violate at least one of Kolmogorov's standard axioms. The earliest example is due to Dirac (1942); similar ideas are present in Feynman (1987), Suppes and Zanotti (1991), Khrennikov (1995, 1997), and Kronz (2007). For a survey, see Mückenheim et al. (1986). Thus there is reason to question whether a qbit can be assigned a probability that satisfies all of the Kolmogorovian axioms and is stationary. We discuss this and related issues in Stecher et al. (2006). For our purposes here, it suffices to assume that such a probability exists, because if not, we already have ambiguity as soon as we obtain the qbits.

Section 4

5. The motivation in Calude et al. differs from ours, but their technique is easily adapted to the issue at stake here. In a sequence of papers, Binmore (1992, 2007a,b) pursues a related issue, namely, that Bayesian decision theory requires subjects to solve the halting problem.

6. We converted these to integers from 0 to 9, in keeping with the practice of Section 3. We found data that were inconsistent across replications, as above; however, all the even numbers arose with a higher frequency than all of the odd numbers. This was true across all 100 replications. When we increased ϕ to 0.2, this became even more pronounced. The simulations were done in *Mathematica*.

7. For example, on page 16, Savage writes that he intends to "confine attention to so small a world that the 'Look before you leap' principle [i.e., that of foreseeing every contingency] can be applied there. I am unable to formulate criteria for selecting these small worlds and indeed believe that their selection may be a matter of judgment and experience about which it is impossible to enunciate complete and sharply defined general principles." For more on this point, see Binmore

(2007b).

Section 5

8. Indeed, shuffling has been used in the finance literature to detect the presence of complex autoregressive structures in stock returns. See Scheinkman and LeBaron (1989).

References

N. Baigent. A re-examination of Bossert and Slinko's theory of uncertainty aversion under complete ignorance. Working paper, Institute for Public Economics, Graz University, 2007.

V. D. Barnett. Order statistics estimators of the location of the Cauchy distribution. *Journal of the American Statistical Association*, 61(316):1205–1218, 1966.

T. F. Bewley. Knightian decision theory. Part I. *Decisions in Economics and Finance*, 25(2):79–110, 2002.

K. Binmore. Debayesing game theory. In B. Skryms, editor, *Studies in Logic and the Foundations of Game Theory: Proceedings of the Ninth International Congress of Logic, Methodology and the Philosophy of Science*. Dordrecht, 1992.

K. Binmore. Can knowledge be justified true belief? Working paper, University College, London, 2007a.

K. Binmore. Making decisions in large worlds. Working paper, University College, London, 2007b.

P. Bossaerts, P. Ghirardato, S. Guarnaschelli, and W. Zame. The impact of ambiguity on prices and allocations in competitive financial markets. Working paper, California Institute of Technology, 2007.

W. Bossert and A. Slinko. Relative uncertainty aversion and additively representable set rankings. *International Journal of Economic Theory*, 2(2):105–22, 2006.

C. S. Calude. Algorithmic randomness, quantum physics, and incompleteness. In M. Margenstern, editor, *Proceedings of the Conference "Machines, Computations and Universality" (MCU'2004)*, number 3354 in Lecture Notes in Computer Science, pages 1–17. Springer-Verlag, 2004.

C. S. Calude and K. Svozil. Quantum randomness and value indefiniteness. CDMTCS 291, University of Auckland Centre for Discrete Mathematics and Theoretical Computer Science, 2006.

C. S. Calude, M. J. Dinneen, and C.-K. Shu. Computing a glimpse of randomness. *Experimental Mathematics*, 11(3):361–370, 2002.

G. J. Chaitin. *The Limits of Mathematics*. Springer-Verlag, 1998.

P. A. M. Dirac. The physical interpretation of quantum mechanics. *Proceedings of the Royal Society London (A)*, 180:1–40, 1942.

D. Ellsberg. Risk, ambiguity, and the Savage axioms. *Quarterly Journal of Economics*, 75(4):643–669, 1961.

R. P. Feynman. Negative probability. In F. D. Peat and B. Hiley, editors, *Quantum Implications: Essays in Honour of David Bohm*, pages 235–248. Routledge & Kegan Paul Ltd, London, 1987.

P. C. Fishburn. The axioms and algebra of ambiguity. *Theory and Decision*, 34(2):119–137, 1993.

P. Gärdenfors and N.-E. Sahlin. Decision making with unreliable probabilities. *British Journal of Mathematical and Statistical Psychology*, 36(2):141–153, 1983.

P. Gärdenfors and N.-E. Sahlin. Unreliable probabilities, risk taking, and decision making. *Synthese*, 53(3): 361–386, 1982.

P. Ghirardato, F. Maccheroni, and M. Marinacci. Differentiating ambiguity and ambiguity attitude. *Journal of Economic Theory*, 118(2):133–173, 2004.

I. Gilboa and D. Schmeidler. Updating ambiguous beliefs. *Journal of Economic Theory*, 59(1):33–49, 1993.

I. Gilboa and D. Schmeidler. Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics*, 18(2):141–153, 1989.

T. Hayashi and R. Wada. An experimental analysis of attitude toward imprecise information. Working paper, Department of Economics, University of Texas at Austin and Department of Economics, Keiai University, 2006.

L. Hurwicz. Optimality criteria for decision making under ignorance. Discussion Paper, Statistics 370, Cowles Commission, 1951.

A. Y. Khrennikov. *Non-Archimedean Analysis: Quantum Paradoxes, Dynamical Systems and Biological Models*. Kluwer, 1997.

A. Y. Khrennikov. p-adic probability interpretation of Bell's inequality. *Physics Letters*, 200(3):219–223, 1995.

F. H. Knight. *Risk, Uncertainty, and Profit*. Hart, Schaffner & Marx; Houghton Mifflin Company, Boston, 1921.

B. O. Koopman. The axioms and algebra of intuitive probability. *Annals of Mathematics*, 41(2):269–292, 1940a.

B. O. Koopman. Intuitive probabilities and sequences. *Annals of Mathematics*, 42(1):169–187, 1941.

B. O. Koopman. The bases of probability. *Bulletin of the American Mathematical Society*, 46:763–774, 1940b.

F. Kronz. Non-monotonic probability theory and photon polarization. *Journal of Philosophical Logic*, 36 (4):449–472, 2007.

R. G. Laha. On a class of distribution functions where the quotient follows the Cauchy law. *Transactions of the American Mathematical Society*, 93(2):205–215, 1959.

R. D. Luce and H. Raiffa. *Games and Decisions: Introduction and Critical Survey*. Wiley, 1957.

F. Maccheroni, M. Marinacci, and A. Rustichini. Ambiguity aversion, robustness, and the variational representation of preferences. *Econometrica*, 74(6):1447–1498, 2006.

W. Mückenheim, G. Ludwig, C. Dewdney, P. R. Holland, A. Kyprianidis, J. P. Vigier, N. Cufaro Petroni, M. S. Bartlett, and E. T. Jaynes. A review of extended probabilities. *Physics Reports*, 133(6):337–401, 1986.

K. Nehring. A theory of rational choice under ignorance. *Theory and Decision*, 48(3):205–240, 2000.

K. Nehring. Decision-making in the context of imprecise probabilistic beliefs. Working paper, University of California, Davis, 2007.

E. A. Pazner and D. Schmeidler. Competitive analysis under complete ignorance. *International Economic Review*, 16(1):246–57, 1975.

E. J. G. Pitman and E. J. Williams. Cauchy-distributed functions of Cauchy variates. *The Annals of Mathematical Statistics*, 38(3):916–918, 1967.

A. Rustichini, J. Dickhaut, P. Ghirardato, K. Smith, and J. Pardo. A brain imaging study of the choice procedure. *Games and Economic Behavior*, 52(2):257–282, 2005.

L. J. Savage. *The Foundations of Statistics*. Dover, 1954. Revised republication of 1954 edition, published posthumously in 1972.

J. Scheinkman and B. LeBaron. Nonlinear dynamics and stock returns. *Journal of Business*, 62(3):311–337, 1989.

D. Schmeidler. Subjective probability and expected utility without additivity. *Econometrica*, 57(3):571–587, 1989.

D. Schmeidler. Integral representation without additivity. *Proceedings of the American Mathematical Society*, 97:255–261, 1986.

J. D. Stecher. Subjective information in decision making and communication. In M. Abdellaoui and J. D. Hey, editors, *Advances in Decision Making under Risk and Uncertainty*, Springer Theory and Decision Library. Springer-Verlag, 2007a.

J. D. Stecher. Existence of approximate social welfare. *Social Choice and Welfare*, 2007b.

J. D. Stecher, R. Lunawat, K. Pronin, and J. W. Dickhaut. Decision making and trade without probabilities. Working paper, Norwegian School of Economics and Business Administration and Carlson School of Management, University of Minnesota, 2006.

P. Suppes and M. Zanotti. Existence of hidden variables having only upper probabilities. *Foundations of Physics*, 21(12):1479–1499, 1991.

M.-L. Vierø. Exactly what happens after the Anscombe-Aumann race? Representing preferences in vague environments. Working paper 1094, Queen's University Department of Economics, 2006.

M.-L. Vierø. Contracting in vague environments. Working paper 1106, Queen's University Department of Economics, 2007.

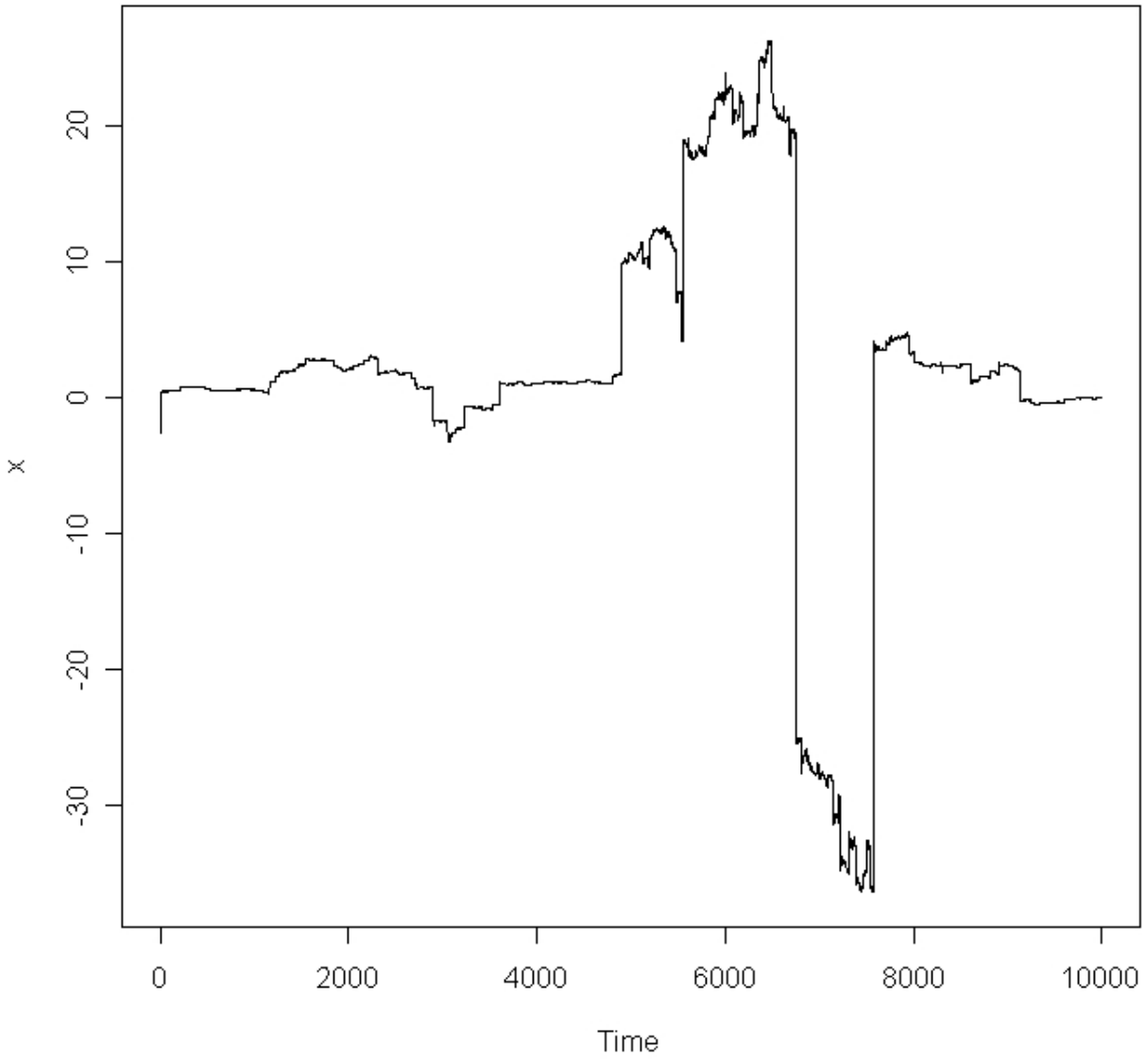


Figure 1: Results of 10,000 ambiguous draws, displayed as a time series

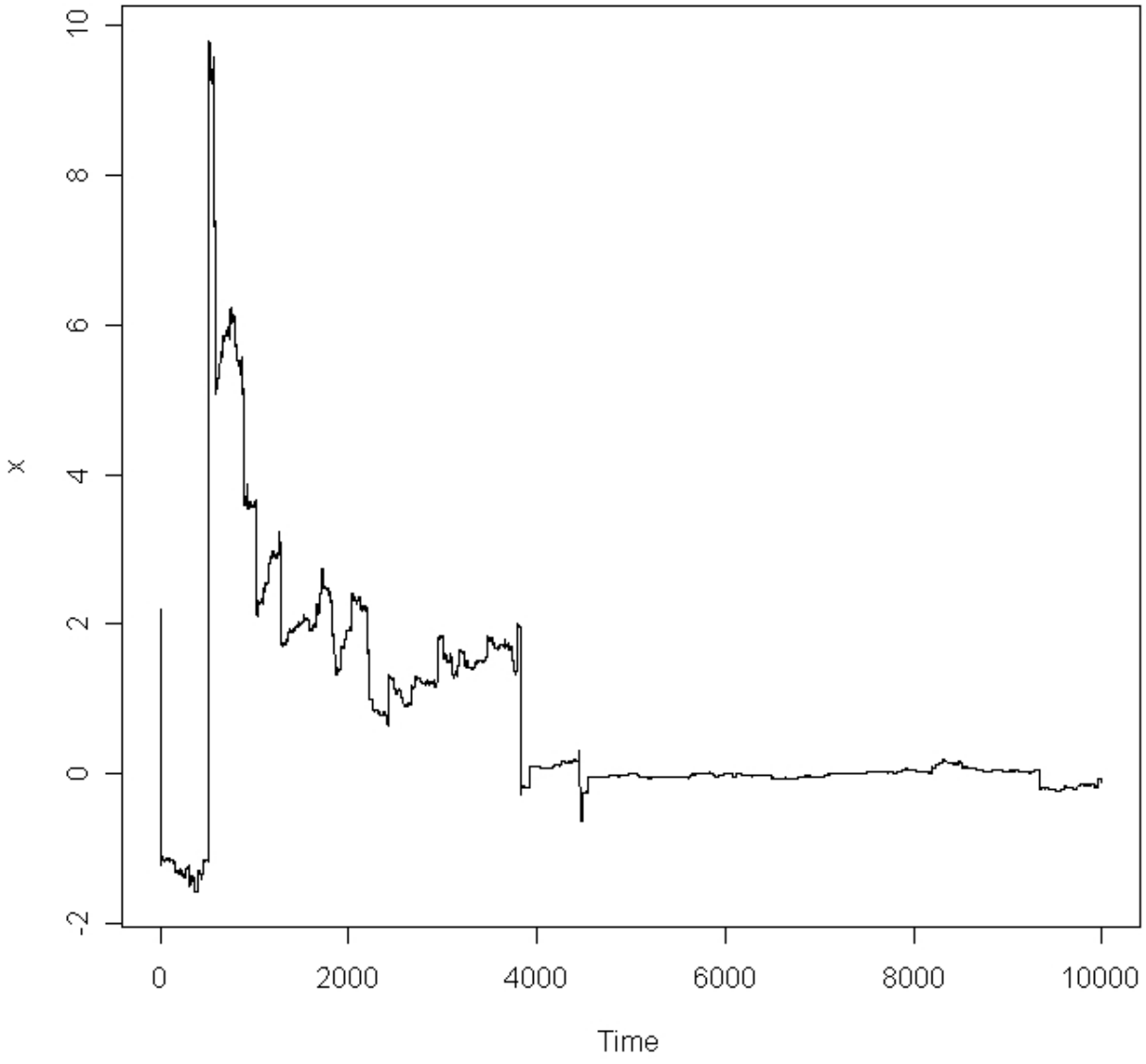


Figure 2: Second replication of 10,000 ambiguous draws

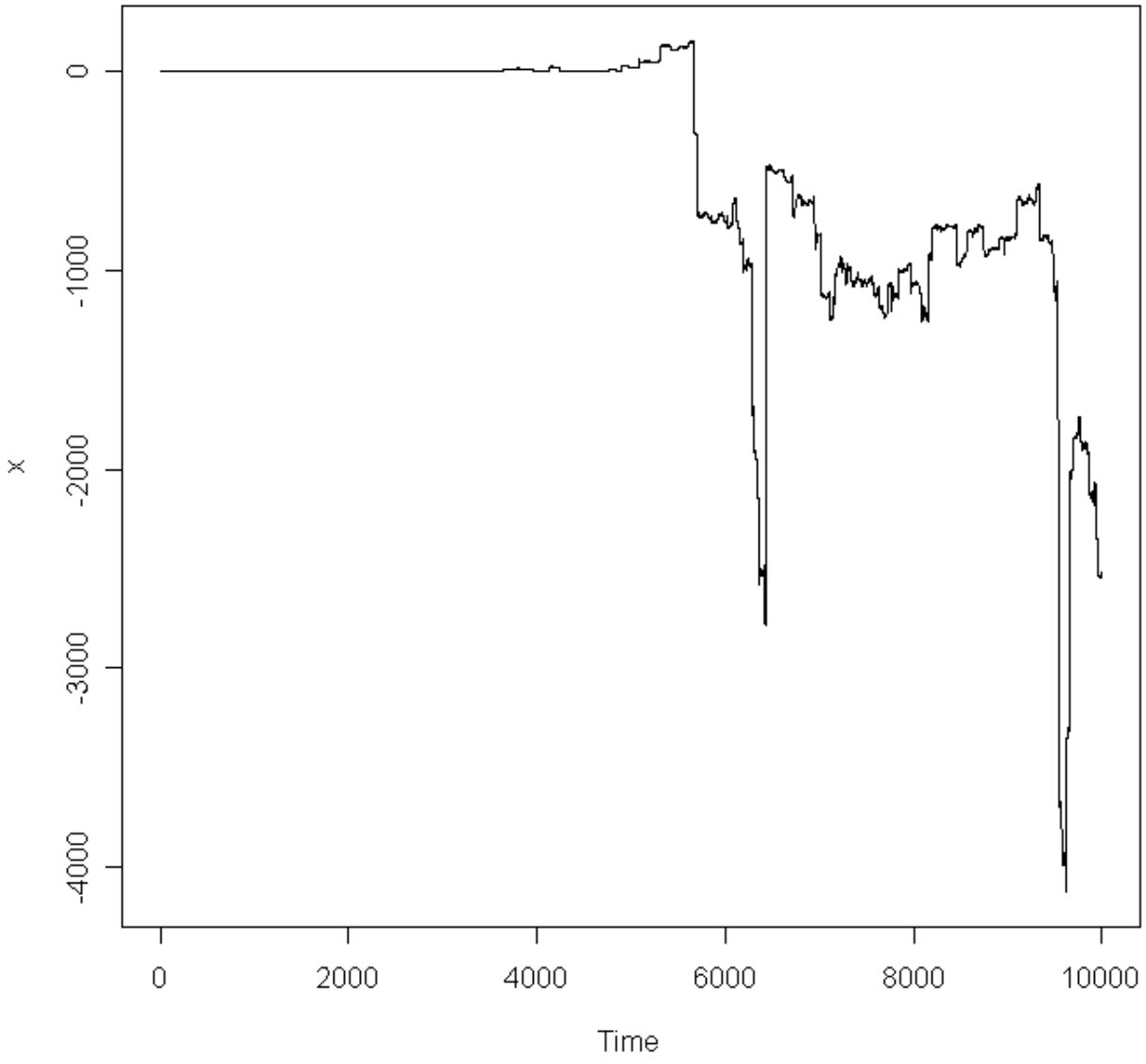


Figure 3: Another replication of 10,000 ambiguous draws

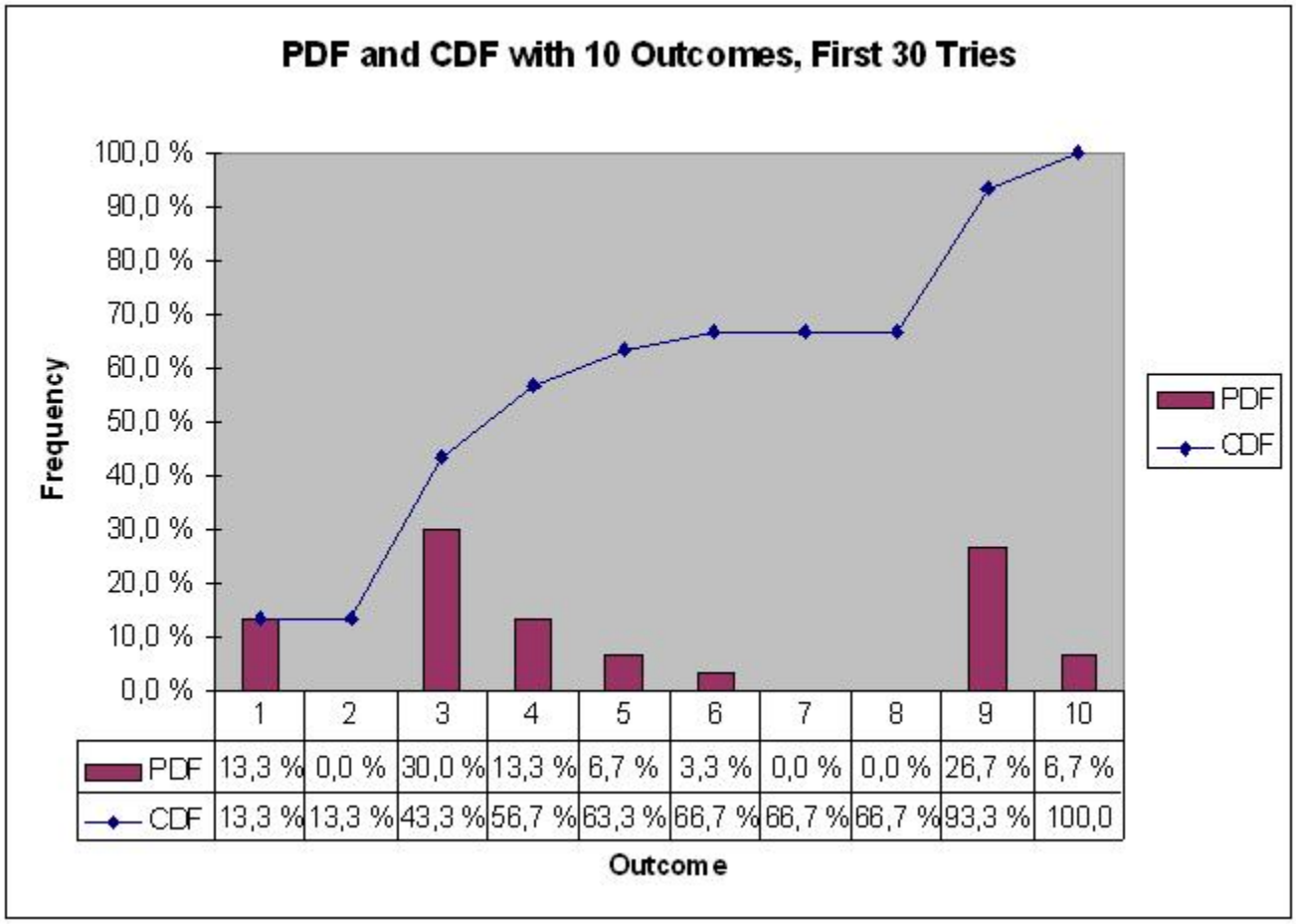


Figure 4: Empirical histogram of 30 ambiguous draws, physical randomization

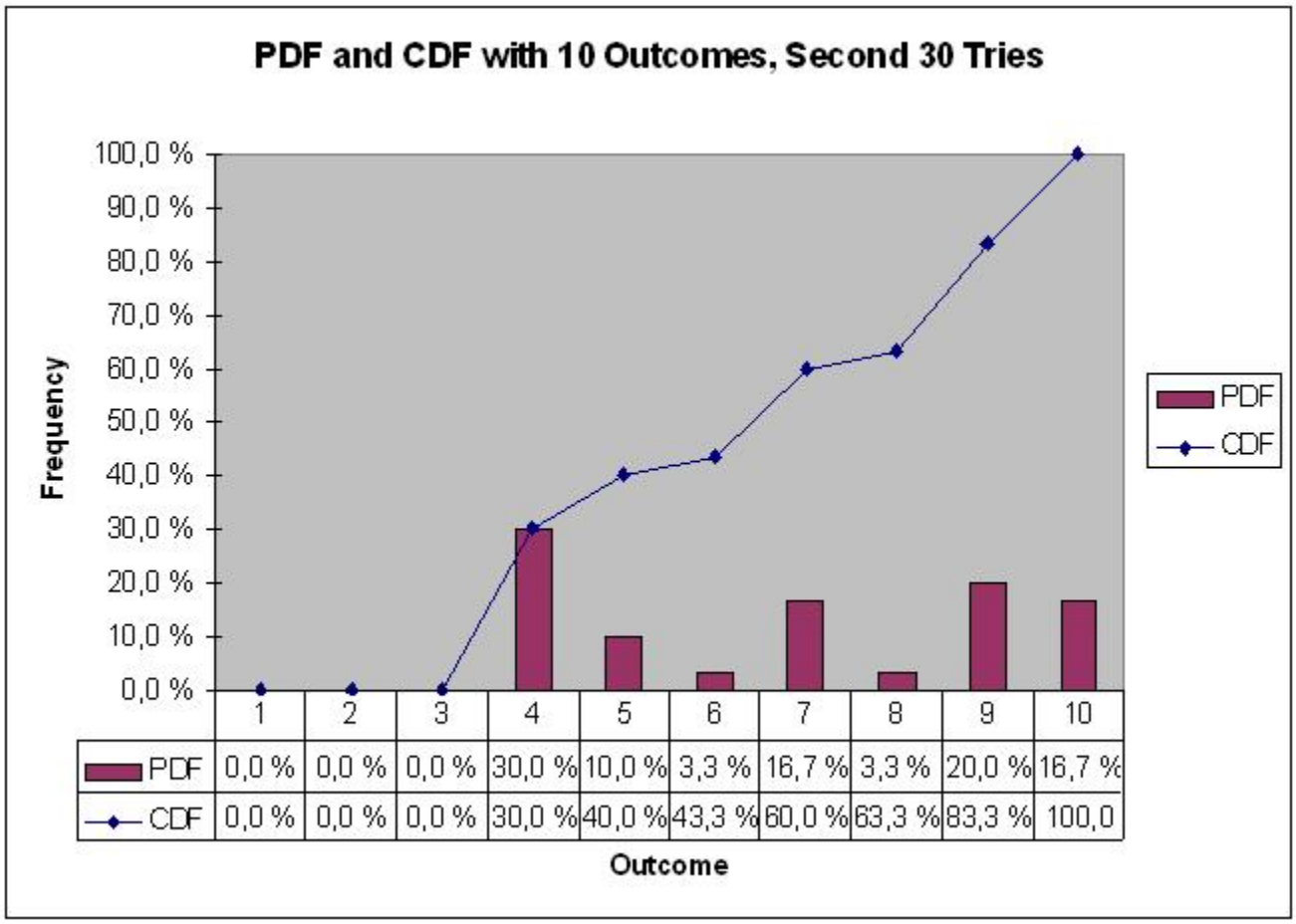


Figure 5: Empirical histogram of 30 ambiguous draws, second replication

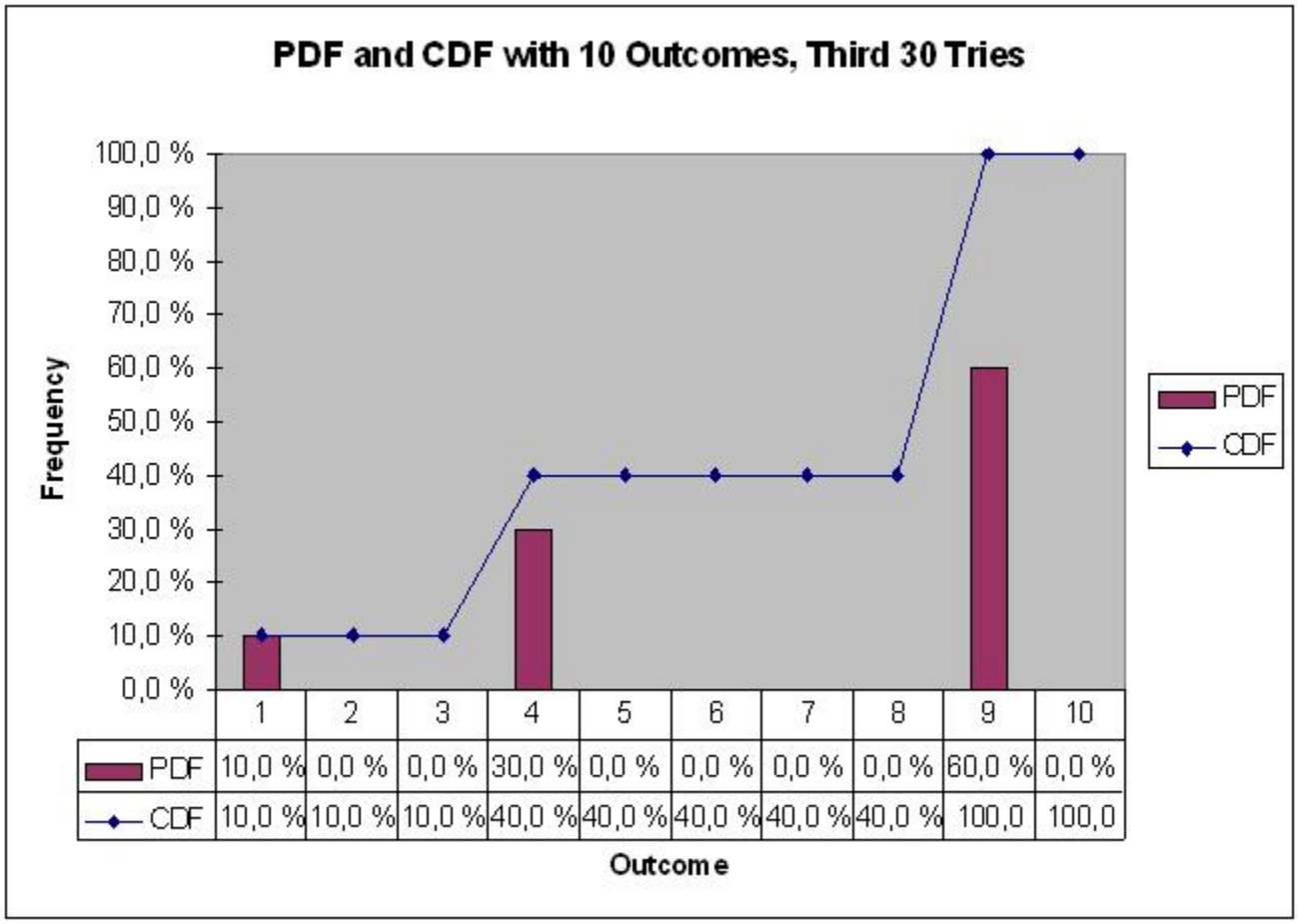


Figure 6: Empirical histogram of 30 ambiguous draws, third replication