An empirical regularity in freshwater sport fishing is that most fish are caught by a small number of anglers (Thompson 1976; Hudgins and Davies 1984; Baccante 1995). Few anglers are thought to have the skill and knowledge to catch a large number of fish (Thompson 1976; McConnell et al. 1995). Food-web structure, gear, and daily bag limits may also contribute to unequal catch distributions (Baccante 1995; McConnell et al. 1995). In a neglected paper, Thompson (1976) proposed an alternative explanation for catch inequality, suggesting that catch inequality is due to random chance. Thompson (1976) used a probabilistic catch model based on the Poisson distribution to show that approximately 10% of anglers can catch approximately 90% of fish by random chance alone. As a consequence, the observation that most fish are caught by a small minority of anglers does not, in the absence of other evidence, establish that skill, equipment, ecology, or policy necessarily determines fish catch (Thompson 1976). Despite this theoretical finding, Thompson (1976) was unable to support his probabilistic catch model with creel survey catch data (his data were significantly different from the Poisson distribution as judged by a chi-square test).

Catch data for angler inequality studies are usually generated by creel surveys that record angler catch and effort for individuals for complete or incomplete trips (e.g., Thompson 1976; Baccante 1995; Cook et al. 2001). Long-term volunteer angler log programs are a low-cost alternative to creel surveys (Cooke et al. 2000; Bray and Schramm 2001; De Jesus et al. 2009). Angler log programs recruit anglers to complete and then return to state fisheries agencies cards or booklets describing trip locations, trip lengths, numbers of fish caught, and types of fish caught. The representativeness of these angler log data to the general population of anglers is uncertain. Bray and Schramm (2001) found that participants in the statewide Mississippi volunteer angler log program had significantly different demographics and fishing preferences from the demographics and fishing preferences of a statewide survey of licensed anglers. However, demographic differences do not necessarily translate to greater or lesser...
angling success. Validation of angler log catch rates with creel survey catch rates has produced inconsistent results. Prentice et al. (1995) found a noisy ($r^2 = 0.42$) but significant positive relationship between catch per unit effort (average number of fish caught per hour) recorded by angler logs and catch per unit effort recorded by state creel surveys in Texas. Bray and Schramm (2001) found an inverse relationship between angler diary catch per hour and creel survey catch per hour for black bass Micropterus spp. in Mississippi but no relationship between angler diary catch per hour and creel survey catch per hour for crappies Pomoxis spp. Despite these uncertainties, long-term angler log programs have been recommended as useful but underexploited data sources for fisheries management (Kerr 2007; De Jesus et al. 2009). Angler log data are particularly well suited for studying remote locations and rarely caught fish (Kerr 2007). Incomplete trip data collected by roving creel surveys do not always capture relationships seen in complete trip data (e.g., VanDeValk et al. 2007). Angler logs record complete trips and thus may be useful for analyses that require complete trip data.

The purpose of this study is to extend and test Thompson’s probabilistic catch model by introducing angling time as a random variable. The probability of catching a certain number of fish is conditioned on time spent fishing. Here I test historic angler log data from six Montana lakes and catch data from a creel survey for deviation from this new probabilistic catch model. I conclude there is no evidence that few anglers catch a large number of fish because of having a better than average aptitude for fishing.

**METHODS**

Data.—I retrieved catch data for lake trout Salvelinus namaycush and largemouth bass Micropterus salmoides from six lakes from the Montana Fisheries Information System (http://fwp.mt.gov/fishing/mFish/). The data were generated by the Montana Fishing Log Program. The program has about 850 volunteer participants (Tipton 2009). Logs are returned quarterly to Montana Fish, Wildlife & Parks. The data consist of total quarterly catch and total quarterly effort (in hours) for each log at each location. Because log data integrate 3-month periods, they can include multiple trips on the same lake. I selected three lakes for lake trout and three lakes for largemouth bass to test for deviation from a random catch model. Lake selection was arbitrary but I included one lake with a small number of log recordings ($n = 19–25$), one lake with a medium number of log recordings ($n = 99–121$), and one lake with a large number of log recordings ($n = 175–202$) for each species (Table 1).

I retrieved 1969 creel survey data from the lower Current River, Missouri, as reported by Thompson (1976) for comparison with the Montana angler log data. I retrieved Gini coefficients for various individual lake creel surveys reported by Baccante (1995) to compare with the Montana angler log data and with the lower Current River creel survey data.

**Analysis of catch inequality.**—For visual assessment of catch inequality (Thompson 1976), I plotted Lorenz curves. The Lorenz curve is a plot of the cumulative proportion of fish caught by the cumulative proportion of anglers. The Lorenz curve is a diagonal line if the catch distribution is equal (i.e., 10% of anglers catch 10% of fish, 20% of anglers catch 20% of fish, and so on). The Lorenz curve is concave if the catch distribution is unequal (e.g., 10% of anglers catch 90% of fish). I calculated Gini coefficients as a quantitative measure of inequality. The Gini coefficient is twice the area between a line of equality (the diagonal) and the Lorenz curve (Cook et al. 2001; Deltas 2003; Kleiber and Kotz 2003). A Gini coefficient of 0 represents perfect equality because the Lorenz curve is diagonal and there is no area between it and the line of equality. The Lorenz curve becomes more concave with increasing inequality. The Gini coefficient increases as the distribution becomes more unequal because the area between the line of equality and the Lorenz curve becomes greater. A Gini coefficient of 1 represents perfect inequality. Gini coefficients reported here are corrected for small sample bias according to Deltas (2003). I also calculated

**TABLE 1.** Measures of the equality of fish catch and goodness-of-fit tests from the statistical analysis. The data are for two species (largemouth bass [LMB] and lake trout [LT]) from six lakes in Montana and the lower Current River in Missouri. The adjusted Gini coefficient is the Gini coefficient corrected for small-sample bias. The goodness-of-fit values for effort are the probability values from the Kolmogorov–Smirnov test applied to time spent fishing; those for the Poisson lognormal distribution (PL) and the Poisson distribution (P) are two-tailed probability values calculated by parametric bootstrapping.

<table>
<thead>
<tr>
<th>Water body</th>
<th>Species</th>
<th>N</th>
<th>Gini coefficient</th>
<th>Adjusted Gini coefficient</th>
<th>Pietra coefficient</th>
<th>PL</th>
<th>Effort</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abbott Lake</td>
<td>LMB</td>
<td>25</td>
<td>0.497</td>
<td>0.518</td>
<td>0.374</td>
<td>0.892</td>
<td>&gt;0.15</td>
<td>0.02</td>
</tr>
<tr>
<td>Echo Lake</td>
<td>LMB</td>
<td>175</td>
<td>0.617</td>
<td>0.621</td>
<td>0.474</td>
<td>0.88</td>
<td>&lt;0.01</td>
<td>0.02</td>
</tr>
<tr>
<td>Lake Mary Ronan</td>
<td>LMB</td>
<td>99</td>
<td>0.459</td>
<td>0.464</td>
<td>0.343</td>
<td>0.972</td>
<td>&gt;0.15</td>
<td>0.02</td>
</tr>
<tr>
<td>Upper Stillwater Lake</td>
<td>LT</td>
<td>19</td>
<td>0.249</td>
<td>0.263</td>
<td>0.234</td>
<td>0.284</td>
<td>&gt;0.15</td>
<td>0.08</td>
</tr>
<tr>
<td>Fort Peck Lake</td>
<td>LT</td>
<td>121</td>
<td>0.537</td>
<td>0.541</td>
<td>0.393</td>
<td>0.968</td>
<td>0.114</td>
<td>0.02</td>
</tr>
<tr>
<td>McGregor Lake</td>
<td>LT</td>
<td>202</td>
<td>0.641</td>
<td>0.644</td>
<td>0.497</td>
<td>0.922</td>
<td>0.116</td>
<td>0.02</td>
</tr>
<tr>
<td>Lower Current River</td>
<td>LT</td>
<td>911</td>
<td>0.708</td>
<td>0.709</td>
<td>0.595</td>
<td>0.814</td>
<td>0.02</td>
<td></td>
</tr>
</tbody>
</table>

*Although this is significantly different from the lognormal distribution, I observed qualitatively that the distribution of effort still exhibits good fit on a lognormal probability plot (not shown).
Pietra coefficients as measures of inequality (Kleiber and Kotz 2003). The Pietra coefficient is interpreted as the proportion of fish caught that would have to be reallocated among anglers in order to achieve a uniform catch distribution. Mathematically, the Pietra coefficient \( (P_c) \) is the relative mean deviation of the Lorenz curve. The Pietra coefficient is calculated as

\[
P_c = 0.5 \sum_{i=1}^{n} c_i - \frac{1}{n},
\]

where \( c_i \) is the proportion of total catch obtained by angler \( i \) and \( n \) is the number of anglers in the sample. The Pietra coefficient ranges from 0 to 1. A Pietra coefficient of 0.5 indicates that 50% of fish caught by a group of anglers would need to be redistributed among the anglers to achieve an equal distribution.

**Random catch model.**—Consider the scenario where anglers encounter fish at a constant rate. There are two possible outcomes to each encounter: catch or no catch. If the probability of making a catch remains low and constant, the catch distribution will approach the Poisson distribution as the number of encounters in the sequence approaches infinity (Thompson 1976). The mean catch size is the product of time spent fishing and an unknown encounter intensity constant representing the number of encounters per hour (Thompson 1976; McConnell et al. 1995).

Anglers accrue fishing time over the course of a fishing season. Not all anglers spend the same amount of time fishing. Consequently, time spent fishing can be represented as a probability distribution, not a constant (Thompson 1976). If a series of random events determines the growth of total fishing time independently of the total amount of fishing time already accrued, then distribution of fishing time for a large number of anglers will converge to the lognormal distribution as a result of the Central Limit Theorem (Lewontin and Cohen 1969; Kleiber and Kotz 2003; Mitzenmacher 2004). Because the mean of the Poisson distribution is the product of the lognormal distribution of time spent fishing and the unknown intensity value, fish catch should follow a compound Poisson–lognormal distribution if angler success is determined by chance alone (Bulmer 1974). Johnson et al. (2005) provide a detailed description of the Poisson–lognormal distribution.

I calculated the maximum likelihood fit of the Poisson–lognormal distribution for the distribution of quarterly catch data for each lake and the lower Current River according to Bulmer (1974) with the poilogMLE package for R (http://www.r-project.org/). I tested goodness-of-fit by comparing the log-likelihood of the fitted distributions to the log-likelihoods of parametric bootstrapped distributions \( n = 1,000 \) (Engen et al. 2002; Johnson and Omland 2004; Rayner et al. 2009). Deviation is expected if processes other than chance affect the fish catch distributions. I also tested for goodness-of-fit to the ordinary Poisson distribution proposed by Thompson (1976), using maximum likelihood estimation and parametric bootstrapping \( n = 100 \). I tested goodness-of-fit of effort to the lognormal distribution using the Kolmogorov–Smirnov test.

Effort was not provided for the creel survey catches and hence was not tested for deviation from the lognormal distribution.

**RESULTS**

There were large areas between the line of equality and Lorenz curves for all lakes, indicating highly unequal fish catches (Figure 1). The symmetrical shape of the Lorenz curves is indicative of a lognormal type distribution (Kleiber and Kotz 2003). Upper Stillwater Lake is an exception: its nonsymmetrical shape is more characteristic of a Pareto distribution (Kleiber and Kotz 2003). The area between the line of equality and the Lorenz curve is smaller for Upper Stillwater Lake than for the other lakes (Figure 1). Upper Stillwater Lake has a Gini coefficient of 0.26, lower than that of the other lakes, which range from 0.46 to 0.64 (Table 1). The Pietra coefficients indicate that between 23% and 50% of fish would have to be reallocated among anglers to create an equal catch distribution. With the Lorenz plots, these Gini and Pietra coefficients indicate considerable catch inequality in all six lakes.

Inequality in the lower Current River creel survey (Gini coefficient = 0.71; Pietra coefficient = 0.6) is more extreme than inequality in the Montana angler logs (compare Figures 1 and 2). Nearly 60% of anglers did not catch any fish, and approximately 20% of anglers caught 80% of fish during the survey (Thompson 1976). The shape of the lower Current River creel survey Lorenz curve (Figure 2) is very similar to other published Lorenz curves of creel survey data (e.g., Cook et al. 2001). The striking difference between angler log Lorenz curves and creel survey Lorenz curves reflects the abundance of zero counts in creel survey curves and the absence of zero counts in angler logs. Gini coefficients for the Montana angler log data (mean = 0.51) are considerably lower (more equal) than the Gini coefficients reported for creel surveys (mean = 0.85) by Baccante (1995).

Angler log catches were not significantly different from the Poisson–lognormal distribution in any lake (Table 1). The number of hours spent fishing was not significantly different from the lognormal distribution expected by the random catch model at five of six lakes (Table 1). At Echo Lake time spent fishing deviated significantly from the lognormal distribution; however, I observed qualitatively that the data were well described by the lognormal distribution based on good fit on a probability plot (not shown). Catch was significantly different \( (P = 0.02) \) from the ordinary Poisson distribution for all lakes except Upper Stillwater Lake, where deviation was only nearly significant \( (P = 0.08) \). Low catch logs were fewer and high catches were more frequent than expected from a Poisson process alone. Further, the mean and variance of the distributions were not equal. The overdispersion of catches relative to the Poisson process is consistent with the results of Thompson (1976). Data from the lower Current River creel survey were also not significantly different from the Poisson–lognormal distribution \( (P = 0.814) \) but were significantly different from the ordinary Poisson distribution \( (P = 0.02) \).
FIGURE 1. Lorenz curves (dashed lines) of fish catch for the six study lakes. The diagonal black lines are lines of equality. The closer a Lorenz curve is to the line of equality, the more equal is the distribution of fish catch between anglers. The symmetrical shape of most of the lakes (the exception is Upper Stillwater Lake) is diagnostic of a lognormal distribution. See Thompson (1976), Baccante (1995), or Cook et al. (2001) for examples of Lorenz curves in fisheries research.

DISCUSSION

The angler log data demonstrate excellent fit to the random catch model, consistent with a stochastic generating process. The difference in form and magnitude of inequality between the angler log data and the creel survey data are striking. Creel surveys record catch from single trips. It is likely that an angler will not catch a fish on a single trip. As the amount of time spent fishing increases, the probability of not catching fish declines because the probability of experiencing only one outcome of a sequence of binomial events declines as the sequence gets
The Gini coefficients for the Montana angler log data are lower than the Gini coefficients reported in the literature for creel surveys, suggesting that cumulative catch over time is more equal than catch from single trips. However, this analysis cannot exclude the possibility of avidity bias or reporting bias in angler log data (Cooke et al. 2000). Alternatively, a lack of zero counts in angler log data may not necessarily indicate that angler log program participants are more skilled anglers than the general population as judged by creel survey. Because the angler log records in this study integrate several fishing trips, the record probably includes more encounters; an angler is thus less likely not to catch a fish than when on a single fishing trip. For these probabilistic reasons, we expect differently shaped catch distributions, even if the distributions have the same amount of inequality as measured by the Gini coefficient. This is illustrated by comparing Lorenz curves for the McGregor Lake and the Lower Current River. These records both contain substantial inequality, but the source of inequality in the Lower Current River creel survey, as judged by the location of the point on the Lorenz curve that is parallel to the line of equality, is an abundance of zeros rather than an extreme spread of catch numbers (seen in McGregor Lake).

My analysis focuses on deviation in catches from a random model. I did not test the opposite condition of comparing catch data for deviation from a skill-based model. The concept of skill in recreational fisheries is poorly defined. Leisure science and fisheries management studies do not provide an adequate basis for the construction of a skill-based model because skill is typically a categorical variable self-reported by anglers (see, e.g., Valentine 2004). Because an angler is unlikely to know the distribution of catches and distribution of skill of all other anglers, he or she has no way of adequately judging his or her own skill.

Because the concept of skill is inherently relative, absolute measures such as catch or catch rate may not adequately describe angler skill. Anglers may become less successful with no change in angler skill if the quality of a fishery declines. However, this decline in catch may be captured as a regression of talent in a skill-based model with absolute measures of success (e.g., McConnell et al. 1995). Skill and specialization level may increase as anglers accrue experience. However, not all people become more skilled at specialized recreational activities as they spend more time participating (Kuentzel and Heberlein 2006; Scott and Lee 2010). Thus angling experience is unlikely to serve as a suitable surrogate measure of skill. Further complicating analyses, the anglers that do become more skilled and specialized may change motivation and focus on the size or species of fish caught, not on the total number of fish caught making metrics such as total catch or catch per unit effort inadequate surrogates for anger skill (Chipman and Helfrich 1988). Development of an easily measurable, operational definition of skill would be of great benefit to fisheries research and would provide additional means for testing the question posed in this study about whether catches are different from random.

The main contribution of the probabilistic catch model in this paper is extending the concept of time as a random lognormal variable. Generally, sport fisheries data are standardized for angling time and given as catch per unit effort (e.g., Kerr 2007). This is an implicit acknowledgment that angling time is an important driver of fish catch. Time has a controlling influence on fish catch, changing the shape of the distribution of how many fish are caught in a given amount of time. To demonstrate this, I plotted the probability mass functions of Poisson distributions with arbitrary mean values of 5 and 10 in Figure 3. The Poisson
distribution mean value corresponds to the amount of time spent fishing adjusted by an intensity constant (see methods). More time spent fishing results in a higher expected catch for an individual angler. The shape of the distribution governing the probability of catching a given number also depends on time spent fishing. As time spent fishing increases, the distribution of catch around the expected catch for an individual angler shifts from being right skewed to being symmetrical. My results show that while fish catch is highly unequal between anglers, catch distributions do not differ significantly from what is expected; that is, the probabilities associated with time spent fishing are consistent with stochastic generating processes. This inequality results from a variety of factors, including catchability, which may be species specific or reflect the quality of the fishery, and the mean and spread of time spent fishing. This analysis does not reject the notion that skill, equipment, or other factors may play important roles in determining distributions of catches. Like many null hypothesis tests for distributional form, there is no explicit alternative hypothesis (e.g., Stephens 1974) in this analysis; thus, there was no need to develop a hypothetical skill-based model. This analysis cannot rule out the possibility that skill, equipment, and other factors could result in the same distribution that random chance does. This analysis also cannot rule out an interaction between catchability and skill in adjusting the effect of time on fish catch. This analysis does show that chance cannot be ruled out as the cause of highly unequal fish catch distributions. Probability provides an adequate alternative explanation to differences in skill, equipment, or other factors for the highly unequal distribution of freshwater sport fishing catch. The possibility that skill influences these patterns awaits a rigorous test.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation (DEB 0716869 and DEB 0917696) and the Department of Environmental Sciences, University of Virginia. The editor, the associate editor, three anonymous reviewers, Michael Pace, Charles Madenjian, and K. T. Perry provided helpful comments that greatly improved the quality of the manuscript.

REFERENCES