Statistical Estimates of Rare Stamp Populations

David L. Herendeen and Gary C. White

ABSTRACT. This paper describes a statistical method for estimating the population of rare stamps from auction catalogs, price lists, expert certificates, and other generally available records. The method presented was developed by biologists to estimate animal populations. Such estimates are done by first capturing, marking, and releasing specimens and then recapturing them. From these data, statistics may be developed to estimate the total population. The latest-generation computer software used for such analyses, called MARK, was developed by Gary C. White and others at Colorado State University. This paper explains how MARK may be used by everyday philatelists interested in estimating the number of rare or very scarce stamps or covers in their collecting area. The methods described do not require one to be a mathematician to use them successfully. The methodology is then applied to four test cases in order to illustrate the efficacy of the approach.

PHILATELIC BACKGROUND

One of the most difficult tasks facing the stamp collector, researcher, and exhibitor is determining the rarity of particular stamps, whether alone or on cover. This is important for exhibitors because it allows them to make claims as to the rarity of items in their collections in a quantitative manner. This is usually done with statements such as "number reported," "number recorded," "number seen by the exhibitor," or, most important, number according to a recognized expert or group of experts with published results. Generally, such numbers are based on censuses conducted by a specialist, or groups of specialists, often over prolonged periods of time.

However, experience has shown no matter how well a census is done, just as with the U.S. population census, not every item that exists will be included, and other individuals will subsequently come to light. Consider, for example, Alexander's monumental work (Alexander, 2001), in which he identified and counted the covers franked with either or both of the first two U.S. stamps (Scott 1 and 2)[1] owned by more than a hundred collectors over 25 years (Scott Publishing Co., 2012a, 2012b). This census resulted in a count of nearly 13,000 covers at the time of publication. A recent update to this census (Scheuer, 2012) shows at least 1,300 new examples that have been identified in the intervening years.

Since it is neither practical nor possible to exactly count any given stamp population, perhaps a method that can provide a statistically valid quantitative estimator can be developed by transferring an existing technology to the philatelic domain. That is the intent of this paper.

*David L. Herendeen (deceased), Institute for Analytical Philately, Inc., P.O. Box 8035, Holland, Michigan, 49422-8035, USA. **Gary C. White**, Department of Fish, Wildlife, and Conservation Biology, Colorado State University, 1474 Campus Delivery, Fort Collins, Colorado, 80523, USA. Correspondence: Gary.White@ColoState.edu.*
Manuscript received 10 April 2013; accepted 26 April 2013.

STATISTICAL BACKGROUND

A branch of statistics that originated in the field of ecobiology is devoted to developing methods for estimating the population of animals in a specific geographical area. Such techniques, whose origins can be traced to the 1930s, are commonly called capture-recapture methods. For interested readers, the historical background is summarized in White et al. (1982). A brief overview of these methods, intended to address the average philatelist, is presented in the following sections.

The Capture-Recapture Model

The basic premise of this statistical methodology is that by trapping animals and marking them, one can recapture them at a later time and by simply counting the number of recaptured animals can estimate the total size of the animal population. Two important assumptions must hold. The first must be the population is closed, that is, there are no births or deaths (demographic closure), and the second must be there is no movement in or out of the area (geographic closure).

As a simple example, suppose that one wants to estimate the number of rabbits in Sherwood Forest. A number of rabbit traps are set throughout the forest. The traps are checked the next morning. The traps were successful in capturing $n_1 = 32$ rabbits. Each rabbit is then identified with, perhaps, an earmark and released. The traps are then taken away. One week later, the traps are reset. This time, $n_2 = 50$ rabbits are captured. But only one of these was earmarked ($r_2 = 1$). Assuming that the probability of trapping any specific rabbit is equal to that of any other rabbit, the chances of capturing a marked rabbit may be estimated by

$$\frac{r_2}{n_1} = \frac{n_2}{N} = \frac{1}{32}.$$

It is possible to make a simple estimate of the total population,, using the Lincoln-Petersen estimator (Cooch and White, 2012). The result is

$$\hat{N} = \frac{n_1 n_2}{r_2} = \frac{32 \quad 50}{1} = 1{,}600.$$

This is the general idea, but as one may expect, the real problem is not so simple.

Program Mark

Colorado State University has been at the forefront of research in this area since the 1970s. White and others have developed several generations of computer software designed to solve this problem using various levels of complexity. This research has led to the program MARK (http://warnercnr.colostate.edu/~gwhite/mark/), which embodies the majority of proven methods for solving capture-recapture problems (White et al., 1982; White and Burnham, 1999).

As will be discussed later in this paper, the methods most usable to estimate rare stamp populations are those that assume a closed population. As noted earlier, this simply means that there is a stable population: no births or deaths and no immigration or emigration. Readers might observe that there may be several instances of "death" in a stamp population. This would include copies destroyed by mishandling or war, copies that "mysteriously" disappear, and copies that find their way into museum collections. This will be discussed in the next section. Also, there is no movement in or out of the area being sampled. As will be seen, this assumption is always met for the philatelic application.

Program MARK has a number of models for closed population analyses. There are four possible model combinations that consider different population dynamics of animals, but do they have application to stamps?

The four models, first proposed by Otis et al. (1978), are as follows:

- M_0: In this simplest model, it is assumed the probabilities of capture and recapture are constant. This means there is no difference in the chances of capturing or recapturing a specific animal on any occasion.
- M_t: This model allows for a variation of capture and recapture probabilities as a function of time. Each animal may have a different capture probability for different times, but these probabilities will be constant for all animals for each trapping occasion.
- M_b: This model allows for the behavior of specific animals to affect their probability of capture. For example, there are some animals that might enjoy the trapping (trap happy) and those that do not (trap shy). However, the initial capture probability is the same across time. Further, this model works by seeing a decline in the number of new animals captured on succeeding occasions; that is, there is a depletion of animals. When such a decline is not happening, then the result is an unidentifiable estimate.
- M_h: This model considers the heterogeneity of the population. Heterogeneity in this context means the capture probability is different for each animal, but there is no variation across time.

Beyond these models, it is also possible to consider these different factors (M_t, M_b, and M_h) in pairs (e.g., M_{tb}) or even taking all three possibilities simultaneously, resulting in M_{tbh}. Some of these combinations result from the work of other researchers, but a detailed list of these is not necessary for this paper.

The most interesting question is which of the four possible models most accurately reflects a rare stamp population? As will be seen, the estimation of rare stamp populations is quite similar to the animal population estimate. Since the stamps are neither sentient nor able to learn, there can be no behavioral differences.

THE STATISTICAL ASSUMPTIONS

The typical capture-recapture algorithms have broad applicability, but the single assumption that greatly simplifies their application to the stamp estimation problem is the population is closed. While it is true stamps may be lost or destroyed over time, this is usually negligible. On the other hand, it is possible that a large percentage of very rare stamps may be "lost" to museums. Is it necessary to consider this? Yes, it should be considered, and the easiest way to model this is to simply remove such stamps from the study and then add them back into the results. For example, suppose four of ten copies of a rare stamp are in museums. Then an analysis is performed for the other six stamps. Assume the model estimates that there is a population of eight stamps. Then we add the four museum copies (which are known to exist) to the estimate, arriving at 12 stamps $(8 + 4)$.

In order to use the MARK program successfully, there are other necessary assumptions, which, for our problem, are as follows:

- Individual stamps maintain their identification over time. That is, one can always identify a given sample from its perforations, centering, cancel, or context (as on a cover). This is easily met for truly scarce stamps.
- All of the stamps are correctly identified when they appear in an auction, price list, or other public forum. Since we are dealing with rare items, we may assume photographs or scans are always available. If not, the occurrence is ignored.
- Each stamp has a constant and equal probability of being sold during a given capture period. Gates (1991) asserted this assumption may be violated because strong upward price pressure may cause more of the stamps to "come out of hiding." The first author feels that most of the philatelic rarities are not purchased for investment but are rather purchased by dedicated collectors and exhibitors who intend to hold them for long periods.

As an example, the mean holding period for the Labuan stamps, discussed later in this paper, has been more than 24 years. This is real dedication.

SELECTING TEST CASES

In order to test the hypothesis that MARK could be used as an important philatelic tool, a number of test cases were selected and executed and the results were analyzed. To this end, four sample problems were chosen. These analyses were all based on earlier detailed census keeping by various researchers. They have been selected to incorporate a number of different characteristics. The test cases are described below.

- The first test case considers one of the most spectacular errors in U.S. philately: the 24¢ inverted "Jenny" airmail stamp issued in 1918 (Scott C3a). This case is important because the population of these iconic stamps is precisely known.
- The second test case uses the census and provenance data for the 8¢ Labuan inverted frame postage due stamp (Scott J6a) performed by Herendeen (2006). This case provides two sets of data, the census through 2005 and then the extension of the census to 2011. Together, these allow the predictive capability of MARK to be examined.
- The third is based on the detailed records of Hawaiian stamps and covers maintained by Gregory (2012). Specifically, the data used relate to the 5¢ provisional overprint issued in 1853 (Scott 7). This case checks that the estimator works as well for covers as for stamps.
- The fourth example studies a scarce U.S. stamp, the data for which were provided by the Robert A. Siegel Auction Galleries, Inc. Web site. This stamp is the U.S. 1¢ stamp made from coil waste and issued from 1923 to 1926 (Scott 594). Only the used stamps were considered for this study. This case checks results for a less precise data set that has many observations based on expert certificates in addition to auctions. With 488 observations, this is also the largest set of the test data.

If all of these test cases result in reasonable estimates that correlate well with reality, then it may be concluded that MARK will be useful to philatelists.

ASSEMBLING THE DATA

Several steps must be taken to gather the input data for the statistical model. These include acquiring the raw data describing the encounters with specific stamps or covers. The data are then divided into time intervals for analysis. These procedures are described in this section.

Raw Data Acquisition

While often laborious, the assembly of raw data needed to perform the statistical analysis is easily done. Using any application software that allows you to enter and sort data (such as Microsoft Excel), you simply record each appearance of the stamp or cover under consideration. Data sources may include auction catalogues, retail price lists, copies of expert certificates, and any other similar records. It is most important that all the items are illustrated to allow positive identification.

In theory, only two pieces of data are required for each observation, but for completeness, three are recommended as a minimum. In fact, the collector may wish to include many different data items so that retracing this step will never be necessary. The minimal items are as follows:

- The sample identifier. A unique number, or other string of characters, which is assigned to each stamp or cover identified.
- The location where the sample was captured. This could be an auction catalog, price list, expertizing certificate, or any other reliable source.
- The date the sample was seen, offered, or sold. Full dates are preferable for completeness, but just the year is usually sufficient.

Examples of these data are seen in Table 1 for the Labuan inverted frame postage-due stamp.

Selecting the Time Interval

Once the basic data have been assembled, the number of time intervals representing sampling occasions must be selected. This is especially important with rare stamps because the data are much sparser than for the typical animal populations.

In a subsequent section, the results of the statistical analyses for different time steps are presented. It appears from these results that no great variations in solutions occur (i.e., less than a 10% variation). Therefore, the best estimates should probably be those obtained from using a one-year time interval as the "encounter occasions."

APPLYING THE STATISTICAL MODEL

Once the data are assembled, they are entered into an Excel spreadsheet. Then, special software, called POPULATION, is used to generate the data for MARK. This software has been developed by the Institute for Analytical Philately, Inc. (IAP).[2] MARK is then executed and the results are produced. This process is described in the following sections.

MARK Input

The MARK program input is a computer text file that has a Windows file extension of .INP. While it is possible to create such files manually, IAP's software tool POPULATION[3] provides a special utility for use with philatelic data. Instead of creating the .INP file manually, the user creates a Microsoft Excel spreadsheet containing the raw survey data.

A partial history of the Labuan inverted frame postage-due data is shown in Table 1. It contains a unique sample number for each of the stamps seen, the "trap," which is the source of the data, an unused column labeled "year," and the exact date that the stamp was seen.

As will be seen in Herendeen (2012) (hereinafter the *Users Guide*), there are many different options available within POPULATION. For example, when executed, POPULATION converts the exact dates into the years for further processing. The user might have just as easily entered only the years into the spreadsheet.

POPULATION is then executed. The output is simply the MARK .INP file. An extract of this file is shown in Table 2. Again, a detailed description of the meaning of this data may be found in the *Users Guide*.

MARK Output

After preparing the input data, the MARK program is executed. Step-by-step details are found in the *Users Guide*. The results of MARK are the statistical computation of the stamp

TABLE 1. The history of the 8¢ Labuan inverted frame postage due stamp.

Sample	"Trap"	Date
1	Paris	1 May 1907
12	Royal Collection	1 Jan 1920
13	Ferrary	26 Apr 1923
17	Harmer-UK	14 Mar 1932
21	Harmer-UK	23 Mar 1935
27	Harmer-UK	16 Dec 1935
28	Harmer-UK	16 Dec 1935
15	BPA	1 Jan 1936
31	RPSL	1 Jan 1936
18	Harmer Rooke	1 Dec 1937
26	Harmer-UK	10 Oct 1938
14	Klein	21 Feb 1939
7	Harmer-UK	21 Oct 1947
7	Harmer-UK	25 Oct 1949
13	Harmer-UK	10 Jan 1950
3	Harmer	14 Nov 1950
22	Friedl	9 May 1951
10	Harmer-NY	9 Apr 1957
5	Harmer-NY	30 Sep 1958
9	Corinphila	18 Nov 1959
29	Siegel	24 Feb 1966
21	Gibbons	13 Feb 1969
1	Robson Lowe	30 Jun 1971
18	Harmer-NY	25 Feb 1975
6	Harmer-NY	17 Oct 1979
24	Harmer-NY	25 Mar 1981
27	BPA	1 May 1981
2	Gibbons	4 Mar 1982
3	Sotheby	6 Sep 1984
16	Christies	23 Oct 1984
32	RPSL	1 Jan 1989
5	Manning	19 Nov 1989
13	Siegel	1 May 1990
13	Western	8 Dec 1990
3	Shreve	7 Nov 1991
8	Robson Lowe	17 Dec 1991
16	Christies	5 Mar 1992
18	Shreve	23 May 1992
5	Holtz	1 Apr 1993
22	BPA	1 Jun 1993
4	Cherrystone	9 Aug 1993
25	RPSL	1 Jan 1994
19	Christies	9 Mar 1994
3	BPA	1 Feb 1995
19	Harmer-UK	18 Dec 1996
1	RPSL	1 Jan 1997
10	Ivy-Mader	1 Jun 1997
17	BPA	1 May 1999
23	Gibbons	15 Oct 1999
9	Singer	1 May 2000
21	Spink	23 Jun 2001
17	Bennett	24 Jun 2001
19	Spink	23 Feb 2003
20	Spink	21 Jun 2003
15	Dealer Stock	1 Jul 2003
19	Dealer Stock	1 Jul 2003
11	BPA	1 Aug 2003
25	Spink	16 Oct 2003
9	BPA	1 Nov 2003
15	Grosvenor	19 Nov 2003
30	Roumet	27 Jan 2004
33	Brun	1 Jan 2005
28	Victoria	7 May 2005

TABLE 2. Extract from MARK data created from Excel spreadsheet.

```
/* Labuan invert  through 2005,  annual encounters */
/* Number Occurrences = 99:  Starting in 1907 ending in 2005 */
```

```
/* 1:3 */ 100000000000000000000000000000000000000000000000000000000000000010000000000000000000000000100000000 1 ;
/* 2:1 */ 000000000000000000000000000000000000000000000000000000000000000000000000000100000000000000000000000 1 ;
/* 3:4 */ 000000000000000000000000000000000000000000010000000000000000000000000000000000000100000010001000000000 1 ;
/* 4:1 */ 000000000000000000000000000000000000000000000000000000000000000000000000000000000000001000000000000 1 ;
/* 5:3 */ 000000000000000000000000000000000000000000000000000100000000000000000000000000000010001000000000000 1 ;
…
…
/* 29:1 */ 000000000000000000000000000000000000000000000000000000000001000000000000000000000000000000000000000 1 ;
/* 30:1 */ 000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000010 1 ;
/* 31:1 */ 000000000000000000000000000001000000000000000000000000000000000000000000000000000000000000000000000 1 ;
/* 32:1 */ 000000000000000000000000000000000000000000000000000000000000000000000000000000000010000000000000000 1 ;
/* 33:1 */ 000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000001 1 ;
/* Number Observed = 33 */
```

TABLE 3. Typical output from program MARK.

	Labuan 99 sampling occasions Real function parameters of {closed capture}			
			95% confidence interval	
Parameter	Estimate	Standard error	Lower	Upper
1:p	0.0139593	0.0024567	0.0098795	0.0196905
2:N	43.416165	5.2347971	37.109245	59.403024

population as well as other statistical estimates. A sample output is shown in Table 3, which provides the summary of the statistical quantities estimated.

For this simple model, these are p, the probability of capture or recapture for all of the stamps on each occasion, and N, the estimate for the total stamp population. These statistics include the estimate of each parameter and the lower and upper bounds of the 95% confidence interval for the estimates. This can be interpreted as the probability of the true answer lying between the lower and upper bound being 95% (although the actual statistical description is more complex).

EXPERIMENTAL VALIDATION

To test the applicability for philatelic analysis of the statistical models used by MARK, four sample problems were assembled. These test cases, along with the purpose for each, were described in an earlier section. The results of these cases follow.

The Inverted Jenny—A Closed-Form Comparison: Test Case C3a

A single sheet of 100 of the U.S. 24¢ "Jenny" airmail stamp was printed in 1918 with the center inverted, as seen in Figure 1 (Scott C3a).

Amick (1986) presents a very detailed study of the appearances of all 96 of the recorded examples and recounts the details of this fascinating discovery and the many intrigues associated with it. No information about the missing four stamps has appeared since the initial dispersal of the stamps.

This test case is particularly interesting for two reasons. First, it allows the fidelity of the MARK program to be determined relative to an exact population size known a priori, and second, this problem was solved in the philatelic setting (Gates, 1992) using an earlier generation of capture-recapture software called CAPTURE (White et al., 1978).

The Amick (1986) data were processed using POPULATION, and the results were used as input to the MARK program.

FIGURE 1. The "inverted Jenny" of 1918. Courtesy Smithsonian National Postal Museum.

FIGURE 2. The Labuan inverted frame postage-due stamp. Courtesy of the Smithsonian National Postal Museum.

TABLE 4. Comparison of statistics for the inverted "Jenny" population.

		95% Confidence Interval	
Study	Estimate	Lower	Upper
MARK	96	96	102
CAPTURE (Gates, 1992)	107	94	121

In the 62 years between 1925 and 1986, there were 427 observations of this stamp. The "closed capture" model was used, and it was assumed that the probabilities of both capture and recapture were equal for all 96 stamps. MARK then found a solution. The results, along with those of Gates (1992), are shown in Table 4.

It is noted that the MARK results are of very high fidelity and predict that there are 96 examples. The upper bound of the 95% confidence interval estimates 102 stamps. The much less precise results from CAPTURE occur simply because the older software had more limitations, especially in the number of occurrences allowed. As a result, Gates (1992) had to make certain assumptions that resulted in much greater variations from the known number of stamps.

The results shown in Table 4 indicate the methodology embodied in MARK is well suited to solving the philatelic population problem. For the remainder of the paper, the notation "[lower CI,prediction,upper CI]" will be used to express the statistics. Thus, the results shown in Table 4 may be stated as

MARK: [96,96,102]
CAPTURE: [94,107,121].

The Labuan Postage-Due: Test Cases Labxx

The Herendeen (2006) study identified 33 examples of a variety of the postage-due stamps issued by the British colony of Labuan in 1901 (Scott J6a). These stamps are especially rare because, as shown in Figure 2, they exhibit the inverted frame error. Unlike the inverted airmail stamp discussed above, these stamps were separated into singles prior to sale. They were also canceled to order[4] at the point of sale with a seven-bar elliptical obliterator. The first recorded public appearance was in 1907 at an exhibition in Paris. The study found 63 reported instances of these stamps from 1997 through 2005, whether appearing in public auctions, submitted to expertizing organizations, or found in dealers' stock. There are undoubtedly other appearances not uncovered during the study. Interestingly, the MARK algorithms are known to be unbiased by the fact that some data are missing.

The 63 observations in 99 years represent a very sparse set of data in the statistical sense. If, for example, the statistical model assumes 99 one-year "trapping periods," then only 42 of these periods result in finding a "trapped" stamp.

To determine whether the discretization of time plays an important role in the statistical analysis, three cases were run:

- 10-year time steps (10 occasions),
- 5-year time steps (20 occasions), and
- 1-year time steps (99 occasions).

Table 5 compares the results of these three models. There is no meaningful difference between them, with the population

TABLE 5. Comparison of statistics for three different sampling intervals.

			95% confidence interval	
Test case	Estimate	Standard error	Lower	Upper
LAB10	44	5.71	37	62
LAB20	46	6.26	38	65
LAB99	43	5.23	37	59

FIGURE 3. Cover showing the use of the Hawaiian 5¢ provisional overprint, Scott 7. Courtesy of the Fred Gregory Collection.

estimate varying from 43 to 46 and the 95% confidence intervals varying from 37 to 65. Thus, the sensitivity of the solution to the time discretization is small. However, since the execution time of the MARK program is so fast for the philatelic application, the first author feels it is easier to simply model the data on an annual basis.

As will be seen in a subsequent section of this paper, this sample problem also allowed a simulated solution for the updated basic data, which shows how the statistics can be used as a predictor.

The Hawaiian Provisional On Cover: Test Case Hawaii

The third sample problem is drawn from a Web site called Post Office in Paradise.[5] The analysis considers the scarce use of the 5¢ provisionally overprinted stamp (Scott 7) on cover as shown in Figure 3. Gregory (2012) has archived only 18 such covers encompassing 66 observations from 1909 until 1999.

A MARK input file was created from these raw data. The input matrix had 18 rows and 91 columns. The program was executed and resulted in the following estimates for [lower CI, prediction, upper CI]: [18,18,24].

The convergence of the estimate to the number observed is explained by the large number of occurrences, 91. This is interpreted to indicate that an additional five examples might be found.

The US 1914 1¢ Coil: Test Case US594

The final example analyzes the data for the U.S. 1¢ stamp made from coil waste and issued from 1923 to 1926 (Scott 594), shown in Figure 4. The Scott catalog indicates that sheets comprised either 70 or 100 stamps. Raw data were obtained from the Robert A. Siegel Auction Galleries, Inc. Web site.[6] In addition to auction records, many stamps were reported only when

FIGURE 4. Example U.S. 1¢ green, rotary, perforation 11 stamp (Scott 594). Courtesy of Robert A. Siegel Auction Galleries, Inc. Sale 976, Lot 2170, 29 September 2009.

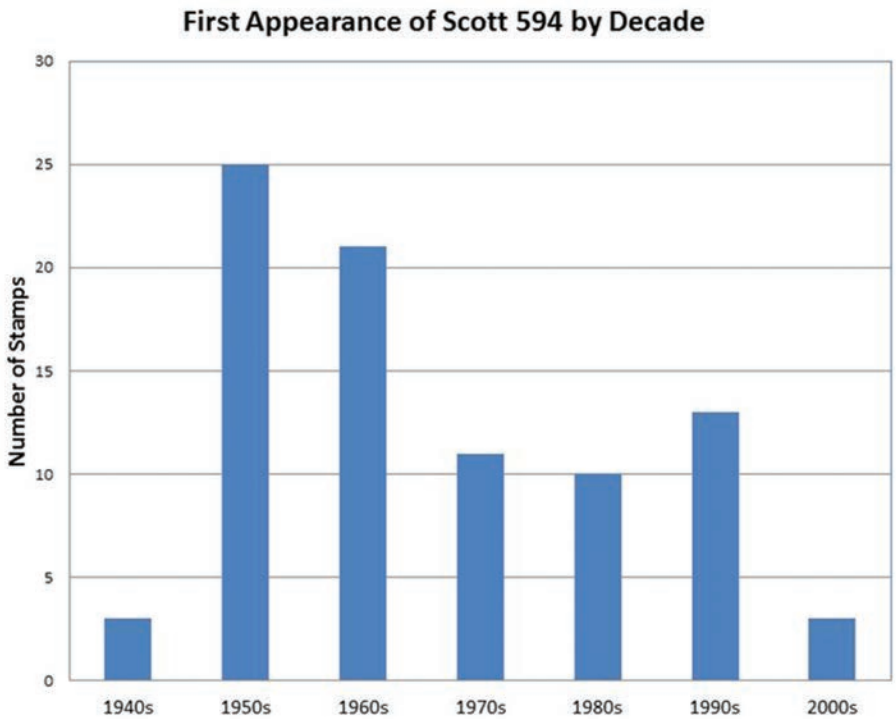

FIGURE 5. First appearance of Scott 594 by decade.

they received certificates of genuineness, most from the Philatelic Foundation.[7] These data were all used to form the observation base. Only the 88 recorded used copies of the stamp were included because they form a larger sample space than the 18 mint examples.

The total number of observations from 1940 to 2008 was 199. A MARK input file was created from these data. The input matrix had 88 rows (one for each stamp) and 69 columns (one for each year). MARK was executed, resulting in [lower CI, prediction, upper CI]: [95,101,115]

This is interpreted to indicate that there are probably another 13 stamps yet to appear and perhaps as many as 27 more stamps. Considering that at least 100, and probably more, of these stamps were printed, this is not surprising.

As an additional check, the plot shown in Figure 5 shows the frequency of appearance by decade. Three new examples appeared in the 2000s and 13 in the 1990s. There is no reason to believe that others will not appear in the coming years.

PREDICTING THE FUTURE

Four test cases have been analyzed using the MARK program. The results appear to be consistent with the actual stamp populations as measured by censuses. Now, how can these statistical estimates be used to predict what might happen in the future? Two methods can help in determining this. The first is to simulate the passage of time by incrementally analyzing a stamp and the second is to actually look at the results of a real case. These are both described in this section.

Test Cases US594

With 199 observation of this rare stamp, the U.S. 1¢ stamp of 1923 affords the best opportunity to do a simulation of the predictive power of these statistical methods. Recall from Figure 5 that new "discoveries" have been made across the seven decades since the first reported sighting of the stamp.

A simulation will now be made assuming that this methodology had been available in 1980. This is done by simply truncating the database of observations to any given year. To that end, the statistics were computed for periods through 1980, 1990, 2000, and 2008 (the end of the available data). The results of these four analyses are shown in Figure 6. This figure shows, for each of the decades, the upper and lower bounds of the 95% confidence interval (UB 95% CI and LB 95% CI) and the estimate of the number of stamps. One of the interesting characteristics of statistical methods is the more items we see, the more we predict exist. This appears to be stabilizing during the full time range. The final estimates [95,101,115] are certainly plausible and would indicate that the number printed would be at least 140 or 200 depending on the sheet size.

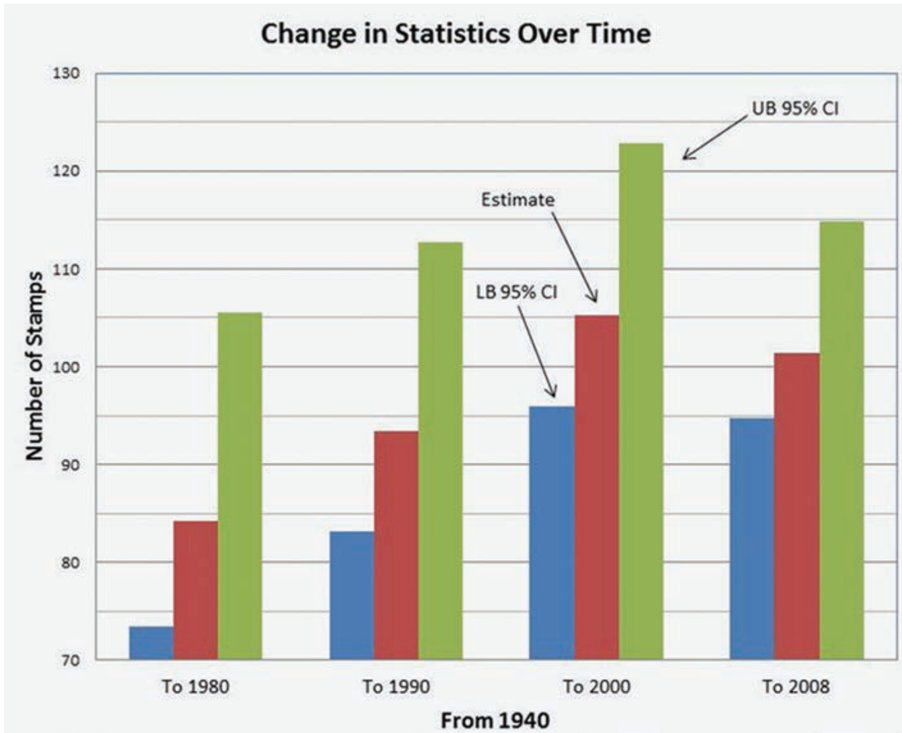

FIGURE 6. Stability of simulated MARK solutions over time for Scott 594. Blue: lower bound, red: estimate, green: upper bound.

Test Case Labnew

As seen previously, the Herendeen study was published in 2006, and the data used in the study ended in 2005. The results of the MARK analysis were [36,41,54]. Therefore, although only 33 stamps were in the census, the model estimated that there would be 41 examples. What has actually happened?

In the intervening seven years, there have been 16 new observations of the Labuan inverts. Most important, four of them were not included in the 2006 census. In other words, there are now 37 known examples. Again, the more items seen, the more probably exist.

The new data were executed by MARK. The results were [40,45,57]. This shows that there is a useful predictive aspect to the statistical analyses.

ADVANCED STATISTICAL MODELS

As noted at the beginning of this paper, other models can be used to improve the estimates for populations of animals. In addition to M_0, which has been used for our sample problems, there are also M_t, M_b, and M_h. Furthermore, these may be applied in pairs (e.g., M_{tb}), or even all three possibilities can be used simultaneously, resulting in M_{tbh}. The multiple-model approach leads to a set of models, each with an estimate of *N*, and hence leads to model selection and how to combine the estimates from multiple models. This is a topic for future work and is beyond the scope of the current paper.

OTHER OBSERVATIONS

Usually, statistical methods are at their best when large samples of data are available. As seen, this is generally not the case for rare stamps, where individual encounters may span decades. How can the MARK program handle situations such as<bl>

- stamps with a known upper bound or
- ultrarare stamps (<4)?

It has already been shown in test case C3A that excellent results were obtained for a stamp whose population has a known upper bound.

A simple test case was created to determine how MARK handles an ultrarare stamp of which only two examples have been seen after 150 years. The subject stamps selected were the U.S. 1¢ Z grill stamps of 1867 (Scott 85A). Again, the Robert A. Siegel Auction Galleries, Inc. maintains an online database of censuses for some rare stamps, including this one. The encounter histories for the two stamps are quite simple. The first example was sold in the mid-1920s (assume 1925) to the New York Public Library, where it remains. The second appeared in 1919, 1957, 1975, 1977, 1986, and 1998.

The input data for MARK thus included seven encounter histories from 1919 to 1998. This model was run and resulted in a population estimate of [2.0,2.0,2.0]. The statistical estimate has converged to the actual census number. This is an excellent result.

CONCLUSION

This paper has described a statistical method for estimating the number of rare stamps and covers from available records such as auction catalogs, price lists, and expertizing organizations. The basic method, well established in the field of wildlife biology and implemented in a computer program called MARK, was exercised using four philatelic test cases. In each test case, the results were completely plausible and correlated well with simulated predictions of future appearances of the rare stamps.

In addition, IAP has made available a Microsoft Excel–based application called POPULATION that allows philatelists to go from census data to input to the MARK program. IAP has also written a user's guide for philatelists to use MARK and POPULATION to get results without being a domain expert in advanced statistical analysis.

ACKNOWLEDGMENTS

The authors acknowledge the following individuals who were kind enough to review this work. We acknowledge Thomas Lera, Winton M. Blount Chair of Research, Smithsonian National Postal Museum; Richard Maisel, New York University; and Scott R. Trepel, president, Robert A. Siegel Auction Galleries, Inc., for their review of this paper and for the useful suggestions that they made.

NOTES

1. Stamps in this paper are referenced by their Scott number as given in *Scott Specialized Catalogue of United States Stamps* (Scott Publishing Co., New York, 2010) or *Standard Postage Stamp Catalogue*, Volume 1, Scott Publishing Co., 2008).

2. The Institute for Analytical Philately, Inc. is a nonprofit corporation that sponsors research that develops methods for applying scientific methods to the solution of philatelic problems. For more information, see www.analyticalphilately.org.

3. The Microsoft Excel–based application POPULATION includes software to convert the simple observation data into the INP file used by MARK. The actual input data are entered into the data spreadsheet. A portion of the test case LAB data is shown in Table 1. For complete details, readers are directed to a separate document (Herendeen, 2012) that provides a comprehensive user's manual for these procedures. The PopUtil spreadsheet and manual are available from IAP by going to www.analyticalphilately.org and first selecting "Free Software," then selecting "Population" and following all of the instructions found there.

4. Canceled-to-order stamps are those that have cancellations applied before they are ever sold to the public. This allows the distributor to set a price not related to the face value of the stamps. The Labuan stamps were sold at the British North Borneo office in London.

5. The Web site www.hawaiianstamps.com, most often referred to as Post Office in Paradise, was created by and is maintained by Fred Gregory.

6. Robert A. Siegel Auction Galleries, Inc. is one of the premier auction houses for U.S. material. In addition, their Web site www.siegelauctions.com includes a number of tools and resources. One of the most important of these is census data for many rare U.S. stamps. This was the source of the raw data for the US594 test case. Stamp illustrated is census number 594-OG-11. http://www.siegelauctions.com/dynamic/census/594/594.pdf.

7. The Philatelic Foundation is the leading expertizing organization in the United States. It keeps extensive records of the stamps that have been viewed. For more information, see http://www.philatelicfoundation.org/.

REFERENCES

Alexander, T. J. 2001. *The United States 1847 Issue—A Cover Census*. Canton, Ohio: U.S. Philatelic Classics Society.

Amick, G. 1986. *The Inverted Jenny*. Sidney, Ohio: Amos Press.

Cooch, E., and G. White, eds. 2012. "Program MARK: A Gentle Introduction." In *Closed Population Capture-Recapture Models*, Chapter 14. Missoula: University of Montana. http://www.phidot.org/software/mark/docs/book/.

Gates, C. E. 1992. Prediction of the Number of Existing Copies of Rare Stamps from Auction Records. *American Philatelist*, 106(1):48–49.

Gregory, F. 2012. *Hawaii Foreign Mail to 1870*, New York: Philatelic Foundation.

Herendeen, D. L. 2006. Labuan 8c Inverted Frame Postage Due Stamp. *London Philatelist*, 114(1):3–25.

———. 2012. *Users Guide for Statistical Analysis of Rare Stamps*, IAP Report 2012-001. Holland, Mich.: Institute for Analytical Philately.

Otis, D. L., K. P. Burnham, G. C. White, and D. R. Anderson. 1978. *Statistical Inference from Capture Data on Closed Animal Populations*. Wildlife Monograph 62. Bethesda, Md.: Wildlife Society.

Scheuer, M. 2012. Updating the Census of 1847 Stamps on Cover. *Chronicle of the U.S. Classic Postal Issues*, 64(4):317–328.

Scott Publishing Co. 2012a. *Scott Specialized Catalogue of United States Stamps*. Sidney, Ohio: Scott Publishing.

———. 2012b. *Standard Postage Stamp Catalogue*. Volume 1. Sidney, Ohio: Scott Publishing.

White, G. C., and K. P. Burnham. 1999. Program MARK: Survival Estimation from Populations of Marked Animals. *Bird Study*, Supplement 1, 46:120–138. http://dx.doi.org/10.1080/00063659909477239.

White, G. C., K. P. Burnham, D. L. Otis, and D. R. Anderson. 1978. *User's Manual for Program Capture*. Logan: Utah State University Press.

White, G. C., D. R. Anderson, K. P. Burnham, and D. L. Otis. 1982. Capture-Recapture and Removal Methods for Sampling Closed Populations. Los Alamos National Laboratory Rep. LA-8787-NERP. Los Alamos, N. M.: Los Alamos National Laboratory.