capstract:2611.00062v1 [math.PR] 20 Nov 2026

Estimating settlement size from the wrapping stratigraphy of fondue sets

Elin Marden
Department of Mathematics, North Ferry Institute, Port Avel, Maine

November 20, 2026

Abstract

We consider a household census based on the number of paper layers surrounding a wedding fondue set at its last transfer. A new wrapper records a change of household. The dedications that governed the transfers need not remain legible. If each household has three prospective recipients, the cumulative stopping hazard at depth t is asymptotically quartic. A small proportion of households with one recipient adds a quadratic term which can dominate the estimate of settlement size. We eliminate this term by comparing the fractions of packages deeper than t and $2t$. The resulting estimator is consistent on the $n^{3/4}$ transfer scale. In artificial deposits from a population of 10,000 households, a single-depth inversion gives 5,825 households at 700 sheets. The two-depth inversion gives 10,029 before sampling error. We give the sampling variance and the conditions under which incomplete packages can be scored.

Keywords: collision time, population size, competing risks, household archaeology

1 The counting unit

Consider an assemblage of fondue services from a settlement whose dwelling floors have been lost. Several services retain a laminated block of tissue paper, folded around the case at successive weddings. A section through the flat face of the block counts wrappings even where the ink on the dedication slips has disappeared. The demographic quantity sought here is the number of households from which the wedding guests were drawn. Our observations consist of two counts of packages exceeding specified depths.

In floor-area estimation, a conversion between occupied area and persons supplies the scale [3]. An estimate from broken cooking vessels requires a discard model, since a household can contribute several vessels to a deposit [4]. The conversion needed for a fondue service is in the other direction. An individual vessel can have supplied a considerable number of households. Wear on the burner begins with its first use, after the period to be estimated. We take the paper section from the broad face of the case, away from the folded ends where a single sheet can contribute several plies. One factory wrapper, if present, is deducted.

We study one particular marriage-gift convention. Each receiving household adds a sheet of tissue to the package and signs a leaf in a dedication register fastened outside the paper. It consults its list of prospective weddings and sends the intact package to the first household absent from the register. A household which has exhausted its list keeps the service and removes the

wrapping as a block. A section remains countable after the last dedication slip has perished. For a deposit preserving a readable register, household names would supply additional observations.

Two features of the observation scheme matter. The sectional count includes the last sheet, added before the final household consults its list. The sample consists of completed wrapping blocks. A service intercepted during circulation has a lower count than the one the model assigns it. In particular, the presence of an intact pot in a case does not by itself establish completion.

2 Law of a completed block

Let $n \ge 4$ be the number of eligible households. For each new service, give each household an ordered list of either one or three distinct other addresses, with probabilities p and $1-p$. Conditional on its length, each list is a uniform ordered sample without replacement from the $n-1$ other households. The lists are independent across households. For inference from several services we also assume independent lists across services, as would be appropriate if each service drew a fresh set of eligible weddings. The initial household may be fixed. Write T for the final number of sheets.

This is a birthday collision with a choice of addresses. For an ordinary first repeat one tests one value at each draw [1]. Here all entries must have been visited to prevent a transfer. Unlike the multiple coincidences of Henze [2], the entries which close a three-address book must themselves be distinct.

Proposition 1 (Section depth)**.** *Put* $S(j) = \Pr(T > j)$ *and* $S(0) = 1$*. For* $1 \le j \le n$*,*

$$S(j) = \prod_{i=1}^{j} (1 - h_i), \qquad h_i = p\,\frac{i-1}{n-1} + (1-p)\,\frac{\binom{i-1}{3}}{\binom{n-1}{3}}. \tag{1}$$

A binomial coefficient with top index less than three is zero in this expression.

Proof. A household can receive the service only once. On its arrival the lists opened so far identify the new holder but leave that holder's own list unexamined. Of the $n-1$ possible addresses, $i-1$ already occur in the register. The chance that every entry is among them is the corresponding ratio of binomial coefficients. Averaging over book length gives the stopping hazard. Multiplying the probabilities of continuing gives Eq. (1). ∎

As a check independent of the product calculation, take four households and $p = 1/3$. Each household then has nine equiprobable books, three of length one and six of length three. Starting at a fixed address and tracing all 9^4 arrays gives 729 blocks of depth two, 1,296 of depth three and 4,536 of depth four. Their frequencies agree with the successive differences of Eq. (1).

3 Cancellation at two depths

Let $b = n^{3/4}$. On this scale a small number of households with a single available wedding can affect both the lower and upper parts of the depth distribution. Write $H(j) = -\log S(j)$ where $S(j) > 0$. The archaeological difficulty is that the proportion of short lists must be inferred from the same material as the size of the settlement. Even a collection that could supply the whole distribution of depths would give a poor single-depth conversion if the short lists were missed.

Lemma 2 (Integrated stopping hazard)**.** *If $p\sqrt{n} \to c < \infty$, then, uniformly on bounded intervals of $x \geq 0$,*

$$H(\lfloor bx \rfloor) \;\longrightarrow\; cx^2/2 + x^4/4. \tag{2}$$

Proof. Summing the two terms of h_i through m gives $p\binom{m}{2}/(n-1)$ and $(1-p)\binom{m}{4}/\binom{n-1}{3}$. This counts choices of two or four indices according to their largest member. For $m = \lfloor bx \rfloor$ the two sums tend uniformly on compact sets to the terms on the right of Eq. (2). The largest h_i on such a set is $O(1/b)$. Thus $\sum_{i \leq m} h_i^2 = o(1)$ there, and replacing h_i by $-\log(1-h_i)$ gives the same limit. ∎

Suppose one were to assign three candidates to every wedding. At a single depth t, inversion of the resulting hazard gives the household count $(t^4/(4H(t)))^{1/3}$. A single-candidate wedding can terminate even a shallow block. Consequently this inversion understates the population. At a paired depth the two contributions scale by different factors. Define

$$\Delta(t) = H(2t) - 4H(t), \qquad N(t) = \left(\frac{3t^4}{\Delta(t)} \right)^{1/3}. \tag{3}$$

Theorem 3 (Household count from paired sections)**.** *Suppose $p\sqrt{n} \to c < \infty$ and $t/n^{3/4} \to x$ for $0 < x < \infty$. Then $N(t)/n \to 1$. If $M = M_n \to \infty$ independent complete packages are available, the same convergence holds in probability for the estimate $\widehat{N}$ obtained by replacing both survival probabilities by sample proportions. The probability that this estimate is defined tends to one.*

Proof. By Lemma 2, the quadratic term at $2x$ is four times its value at x and drops out. The quartic remainder gives $\Delta(t) \to 3x^4 > 0$, while $t^4/n^3 \to x^4$. For either depth the variance of the empirical survival proportion is at most $1/(4M)$. Its expectation tends to a strictly positive limit. The two logarithms therefore converge in probability to their population limits and their difference has positive limit. ∎

An independent pilot deposit can supply a depth when the population scale is unknown. For instance, three quarters of its sample median, rounded down, has the required order if the pilot size tends to infinity. This follows by applying the variance bound in the proof at depths just above and below the unique median of the continuous limiting law. Conditional on the pilot, the main deposit is sampled at fixed cutoffs.

The two counts come from the same packages. For fixed n, t with $0 < S(2t) < S(t) < 1$ and $\Delta(t) > 0$, let $u = S(t)$ and $v = S(2t)$. The indicator for exceeding $2t$ is bounded by the indicator for exceeding t. Their covariance is $v(1-u)$. Applying the delta method to the sample means in $4\log\hat{u} - \log\hat{v}$ gives the sampling variance coefficient

$$V = 8(1/u - 1) + (1/v - 1). \tag{4}$$

More precisely, $\sqrt{M}(\widehat{N}/N(t) - 1)$ converges in law to a centered normal with variance $V/(9\Delta(t)^2)$. This uncertainty concerns the deposit. It does not include the finite-population error $N(t) - n$. A zero upper count or a nonpositive fitted contrast gives an unreported estimate. Both events occur in small deposits.

4 Artificial deposits

We fixed a census of 10,000 households, with a one-percent probability of a single prospective wedding at each address. We used depths of 700 and 1,400 sheets. These cutoffs were set before drawing deposits. The exact survival probabilities from Eq. (1) are 0.738066 and 0.145301. If the short lists are omitted from the analysis, inversion at the two depths gives 5,825 and 7,926 households respectively. The paired estimate is 10,029. The discrepancy between the two single-depth estimates can thus arise in material from one occupation phase.

Figure 1 shows what happens if both sections are moved. The upper tail becomes expensive to sample as the cutoffs increase. In particular the closeness of the paired curve to the census line describes the inversion of population probabilities and gives no assurance that a finite collection contains a package of the required depth.

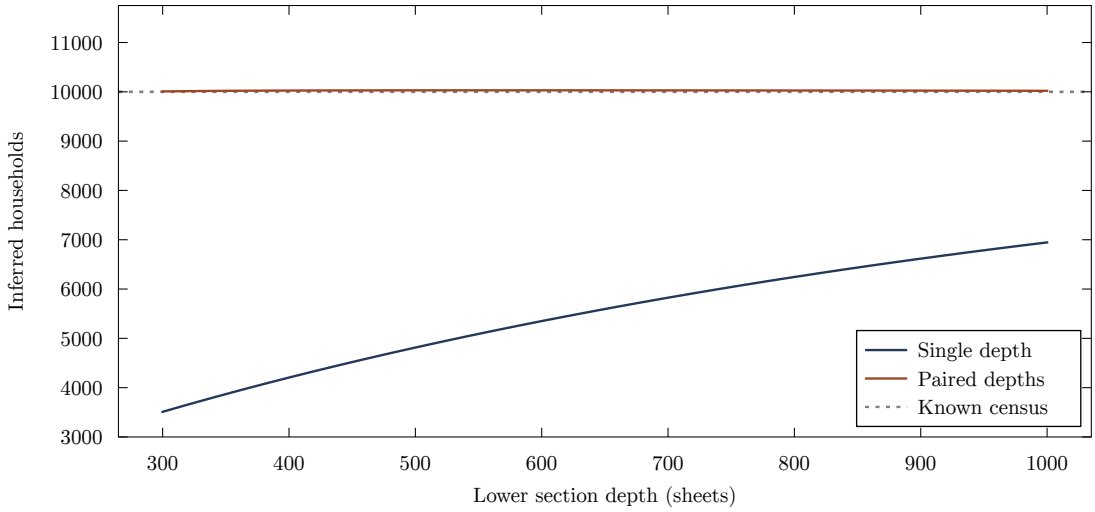


Figure 1: Inferred household count as a function of the lower section depth. The paired calculation also uses twice the indicated depth. Both curves use the exact finite-population probabilities for 10,000 households and a one-address probability of 0.01.

We generated 2,000 independent deposits at each of three sizes. One uniform variate assigns each package to a bin with endpoints 700 and 1,400, using the three probabilities from the finite product. We used Python's Mersenne Twister with seed 47019, processing the deposit sizes in ascending order. The first 128-package deposit had 103 blocks deeper than 700 and 19 deeper than 1,400. Its inferred population was 8,852. Table 1 reports nearest-rank quantiles among deposits yielding an estimate.

Table 1: Recovered household counts in 2,000 simulated deposits per row. The true count is 10,000. Failed inversions are included in the second column and excluded from the quantiles.

Packages	Failed inversions	10th percentile	Median	90th percentile
32	182	7,754	9,488	14,514
128	6	8,705	9,960	12,184
512	0	9,307	10,010	10,929

At 32 packages the positive-contrast restriction removes about nine percent of deposits. The reported median at this size is conditional on being able to issue a census. At 512 packages every deposit furnished both counts and a positive contrast. A larger deposit reduces this sampling variation. It cannot correct the error from treating all lists as long.

5 Incomplete sections and the catchment

Since the estimator requires only two threshold decisions per package, much of a wrapping block can be missing. Suppose a fragment retains 1,401 distinct successive sheets. It contributes to both exceedance counts even if the original depth cannot be recovered. A fragment with 701 sheets and a missing inner face establishes the lower exceedance but leaves the upper one unresolved. Counting this second fragment as a complete package of depth 701 lowers the upper survival estimate and can give a spuriously small settlement.

The effect of unresolved sections can be assessed without assigning a distribution to lost paper. For a known inventory of M packages, let K_i be the number known to exceed cutoff i and let U_i be the number still indeterminate there, for $i = t, 2t$. The remaining packages are known to fall at or below the cutoff. If $K_{2t} > 0$, monotonicity of the logarithm bounds the completed-sample contrast between

$$4\log(K_t/M) - \log((K_{2t}+U_{2t})/M) \quad \text{and} \quad 4\log((K_t+U_t)/M) - \log(K_{2t}/M).$$

The endpoints need not be attainable together with every nesting constraint, but the bounds remain valid. A positive lower endpoint gives a finite upper bound on the inferred count. For the 128-package example, suppose four of the 19 upper exceedances were damaged enough to be indeterminate at 1,400 while remaining positive at 700. Completing the inventory with 15 to 19 upper exceedances gives inferred counts between 8,267 and 8,852 households.

The census refers to households eligible to receive the service. Out-of-town weddings enlarge that catchment beyond the excavated settlement. Reuse of one family’s address book for several services has a different effect. The survival law of a single block can still hold after averaging over books, while the replicate variance calculation fails. Independent packages and a common catchment are thus distinct conditions. Refitting a model for correlated services would require evidence that particular blocks traveled together. The manufacturer’s mark identifies a production lot and by itself supplies no such evidence.

References

[1] Michael Camarri and Jim Pitman. Limit distributions and random trees derived from the birthday problem with unequal probabilities. *Electronic Journal of Probability*, 5:1–18, 2000. doi: 10.1214/EJP.v5-58.

[2] Norbert Henze. A Poisson limit law for a generalized birthday problem. *Statistics and Probability Letters*, 39:333–336, 1998. doi: 10.1016/S0167-7152(98)00076-5.

[3] Raoul Naroll. Floor area and settlement population. *American Antiquity*, 27:587–589, 1962.

[4] Mark D. Varien and James M. Potter. Unpacking the discard equation: Simulating the accumulation of artifacts in the archaeological record. *American Antiquity*, 62:194–213, 1997. doi: 10.2307/282506.