RM - 685

Using Randomized Confidence Limits to Balance Risk – An Application to Medicare Fraud Investigations

Dennis Gilliland
Department of Statistics and Probability, Michigan State University, E. Lansing, MI 48824, USA

Don Edwards
Department of Statistics, University of South Carolina, Columbia, SC 29208, USA

September 6, 2010

Abstract

The National Health Care Anti-Fraud Association (www.nhcaa.org) states that in 2007 over 4 billion health insurance claims were processed in the United States and that fraud amounted to $68 billion. The problem has been featured in a CBS "Sixty Minutes" segment on October 25, 2006 and in a number of NBC News "Fleecing of America" segments, most recently in January, 2010, http://dailynightly.msnbc.msn.com/archive/2010/01/11/2170025.aspx. Additional overpayments come from billings for unnecessary practice and procedures and errors in billings. The recovery of overpayments by Medicare and Medicaid alone is of great national importance.

Guidelines from the Centers for Medicare & Medicaid Services state "In most situations, the lower limit of a one-sided 90 percent confidence interval should be used as the amount of overpayment to be demanded for recovery from the physician or supplier." The *Minimum Sum Method* (Edwards et al (2005)) is based on inverting a family of tests of Hypergeometric distributions to obtain a nonrandomized lower bound for the number of billings among N that are completely in error. This paper shows how to construct a randomized lower bound that serves to increase recovery demands while balancing risks to predetermined levels. The method is connected to the Binomial case by considering the limit as N goes to infinity.

Keywords: randomized confidence intervals, Hypergeometric, Binomial, Medicare overpayment recovery

1. Introduction. Edwards et al (2005) bring attention to Medicare payment populations where payments vary little and a claim is either proper or not, thus making the inherent overcharge (overpayment) either 0 or the entire payment. Medicare examples include populations of payments for motorized wheelchairs and populations of payments for home health care services for periods of a specific length. They propose and investigate a methodology using simple random sampling and any sample size that produces a lower estimate for total overpayment with guaranteed minimum confidence level. Their methodology is based on exact lower estimates for a Hypergeometric, and it is called the *minimum sum method.* Specifically, let M denote the number of population payments completely in error, and suppose a simple random sample of size n finds x payments completely in error. Let L_M denote the lower (1 - α)100% confidence bound for M found by inverting the Hypergeometric test for H_0: $M \leq M_0$. The ordinary minimum sum lower confidence bound for the total population overpayment is then

$$\begin{pmatrix}\text{Total sample}\\ \text{overpayment}\end{pmatrix}+\begin{pmatrix}\text{Sum of the smallest}\\ L_M - x\ \text{unsampled payments}\end{pmatrix}.$$

The methodology guarantees a confidence level of at least (1 - α)100% and does not depend upon a large sample size. Monte Carlo simulations show the advantage that the minimum sum method has over methods based on the Central Limit Theorem (*CLT*) when the population error rate is large, a situation not unusual when there is fraud; in these cases, methods based on the CLT often have achieved confidence level far below the nominal (1 - α)100%. Gilliland and Feng (2010) show how to extend the range of effectiveness of the minimum sum method to populations where payments vary by partitioning payment dollars into packets of uniform size and sampling packets. Edwards et al (2010) study the most extreme choice of packet size, proposing to audit randomly selected pennies.

The recoupment figure set by Medicare in an audit is the lower 90% estimate for total overcharge[1]. The recoupment figure is increased if the 90% lower estimate for the Hypergeometric is increased. This paper shows how to accomplish this through artificial randomization using the same ideas that have been previously applied to develop randomized lower estimates for a Binomial proportion.

Let $B(n, \pi)$ denote the Binomial distribution based on n trials and success probability π; its probability mass function is given by

$$p_\pi(x) := \binom{n}{x}\pi^x(1-\pi)^{n-x},\ x = 0, 1, \ldots, n.$$

Let $H(n, N, M)$ denote the Hypergeometric distribution based on sample size n and a population of size N containing M successes; the probability of x successes in a simple random sample of size n is

$$p_M(x) := \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}},\ x = 0, 1, \ldots, n.$$

Here, and except when necessary for a proof, we suppress the display of dependence of Hypergeometric probabilities on n and N. Nonrandomized and so called *crisp* or *exact*

[1] The CMS (Centers for Medicare & Medicaid Services) prescribes the use of the $\alpha = 0.10$ lower bound as the recovery figure for overpayments by Medicare; we use $\alpha = 0.10$ in our examples.

$(1-\alpha)100\%$ confidence interval estimates for the π in $B(n, \pi)$ or the M in $H(n, N, M)$ have confidence levels in excess of the nominal $(1-\alpha)100\%$. Lehmann (1959, p. 81)[2] considers the $B(n, \pi)$ distributions and shows how the use of artificial randomization allows for the inversion of a family of size α tests that results in randomized interval estimates of π with exact confidence level $(1-\alpha)100\%$. Brown et al (2001, 2002) survey, analyze and compare many interval estimation methods applicable in the Binomial case. Geyer and Meeden (2005) introduce the fuzzy set approach to deal with the problems inherent with crisp interval estimates. Also see Brown et al (2005).

In this paper, we consider the Hypergeometric model and give a construction for a randomized lower estimate of M that enjoys exact confidence level $(1-\alpha)100\%$. The approach is equivalent to inverting the family of size α uniformly most powerful (*UMP*) tests. The test functions are the nonmembership functions in the fuzzy set interpretation developed in Geyer and Meeden (2005).

We consider lower estimates of M, called lower bounds for M, with confidence level $(1-\alpha)100\%$ and $0 < \alpha < 1$. In Section 2, we review the construction of a nonrandomized lower bound for M. In Section 3, as an example we construct the randomized lower bound for M for the example with $\alpha = 0.10$, $n = 5$, $N = 20$. Section 4 gives the construction for the general case, an expression for its expectation and the connection to the Binomial case as $N \rightarrow \infty$. Section 5 contains some applications to illustrate the practical significance of the Hypergeometric case in Medicare payment audits.

---

[2] Lehmann and Romano (2005, pp. 166-167) discuss randomized confidence intervals but do not provide the detail available in Lehmann (1959, p. 81).

Calculations reported in this paper were done with Microsoft Office Excel 2010 or *R* (www.R-project.org).

2. Nonrandomized Confidence Bounds. We fix the population size N and the sample size n with $1 \le n < N$ and refer to the Hypergeometric probability distributions as P_M, $M = 0, 1, \ldots, N$. X denotes a random variable with this distribution. This family of distributions indexed by M has a nondecreasing likelihood ratio in x. Consider the tail probabilities F_M and $\overline{F}_M$ defined by

$$F_M(x) := P_M(X \le x),\, x = 0, 1, \ldots;\, M = 0, 1, \ldots, N \tag{1}$$

$$\overline{F}_M(x) := P_M(X \ge x),\, x = 0, 1, \ldots;\, M = 0, 1, \ldots, N. \tag{2}$$

Consider the integer-valued function

$$L(x) := \min\{m \in \{0,1,\ldots,N\} \mid \overline{F}_m(x) > \alpha\},\;\; x = 0, 1, \ldots, n \tag{3}$$

$$= \min\{m \in \{0,1,\ldots,N\} \mid F_m(x-1) < 1-\alpha\},\;\; x = 0, 1, \ldots, n. \tag{4}$$

The <u>level</u> α nonrandomized *UMP* test of H_0: $M = M_0$ *v.* H_1: $M > M_0$ rejects H_0 if the *p*-value $\overline{F}_{M_0}(x_{obs}) \le \alpha$ and "accepts" (retains) H_0 if $\overline{F}_{M_0}(x_{obs}) > \alpha$, where x_{obs} denotes the observed value of X. (See Lehmann (1959, Theorem 2, p. 68.)) By the stochastic ordering of the family $H(n, N, M)$, if M_0 is rejected then so are all smaller values of M, and, if M_0 is accepted, then so are all larger values of M. The inversion of the family of tests leads to the fact that $\{m \mid L(X) \le m \le N\}$ is a $(1-\alpha)100\%$ confidence set estimate of M. The lower endpoint $L = L(X)$ is a statistic called the *nonrandomized* $(1-\alpha)100\%$ *confidence lower bound* for the parameter M. The coverage probability satisfies

$$C_M(L) := P_M(L(X) \le M) \ge 1-\alpha \text{ for all } M = 0, 1, \ldots, N. \tag{5}$$

(The statistic $U = U(X) := N - L(n - X)$ is the *nonrandomized* $(1 - \alpha)100\%$ *confidence upper bound* for the parameter M.)

Consider the lower bound L for the number M of successes in the population. From definition (3), it is seen to be defined on $\{0, 1, \ldots, n\}$ taking values in $\{0, 1, \ldots, N\}$. Its elementary properties include: $L(0) = 0$, $L(n - 1) < N$, and $L(x) \geq x$, $x = 0, 1, \ldots, n$.

Remark 1. L(x) is a strictly increasing function.

Proof. In this proof we denote the Hypergeometric probability distribution by $P_{n,N,M}$ and use the fact that $P_{n,N,M} = P_{M,N,n}$. Think of drawing a simple random sample of size M from the N by first drawing a simple random sample of size $M - 1$ and then drawing an additional element from the remaining $N - M + 1$. Letting I denote the indicator that this last draw is a success, it follows that

$$\overline{F}_{M,N,n}(x+1) = \overline{F}_{M-1,N,n}(x+1) + P_{M-1,N,n}(X = x)\,Prob(I = 1)$$

$$= \overline{F}_{M-1,N,n}(x) - P_{M-1,N,n}(X = x) + P_{M-1,N,n}(X = x)Prob(I = 1) \leq \overline{F}_{M-1,N,n}(x).$$

Thus, $\overline{F}_{n,N,M}(x+1) \leq \overline{F}_{n,N,M-1}(x)$. Using (3), $L(x) = m$ implies that $\overline{F}_{n,N,m-1}(x) \leq \alpha$ from which $\overline{F}_{n,N,m}(x+1) \leq \alpha$ which implies that $L(x+1) \geq m+1$. □

Later we will have occasion to compare the lower bound $L(X)/N$ for the population rate of success $\pi = M/N$ with the nonrandomized $(1 - \alpha)100\%$ lower bound for Binomial probability π based on X distributed $B(n, \pi)$. From Stapleton (2009, p. 369), the $(1 - \alpha)100\%$ lower bound for Binomial probability π it is

$$L_{binom}(x) = [1/(1+(\nu_1/\nu_2)F_{1-\alpha,\nu_1,\nu_2})] \qquad (6)$$

where $F_{1-\alpha,\nu_1,\nu_2}$ denotes the $1-\alpha$ quantile of the F-distribution with $\nu_1 := 2(n-x+1)$ and $\nu_2 := 2x$ degrees of freedom. If $x = 0$, the lower estimate is taken to be 0.

Example 1. Suppose that $\alpha = 0.1000$, $n = 5$ and $N = 20$. Table 1 provides the nonrandomized lower bound L for M, confidence level 90%. The nonrandomized lower bound for $\pi = M/N$ based on X distributed $B(n, \pi)$ is included for comparison.

Table 1. The Nonrandomized 90% Lower Bound L ($\alpha = 0.1000$, $n = 5$, $N = 20$).

x	0	1	2	3	4	5
$L(x)$	0	1	3	6	10	14
$L(x)/20$	0	.05	.15	.30	.50	.70
$L_{binom}(x)$	0	0.021	0.112	0.247	0.416	0.631

Figure 1 is a plot of the coverage probability $C_M(L)$. Figure 2 is a plot of the expected value of the nonrandomized lower estimate $L(X)$, which we denote as $E_M(L)$.

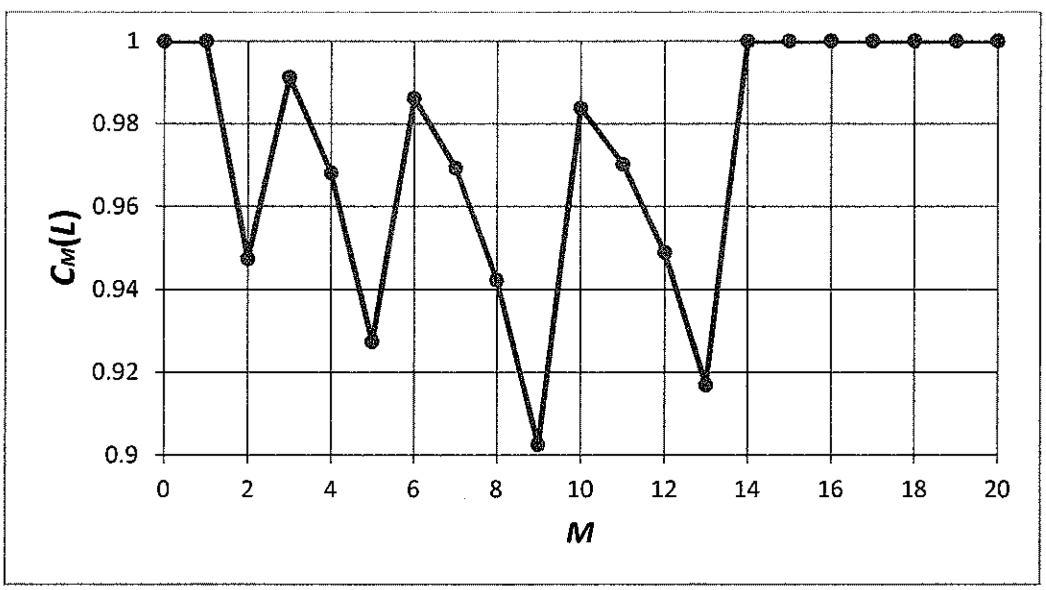

Figure 1. Coverage Probability of the Nonrandomized 90% Lower Bound $L(X)$.

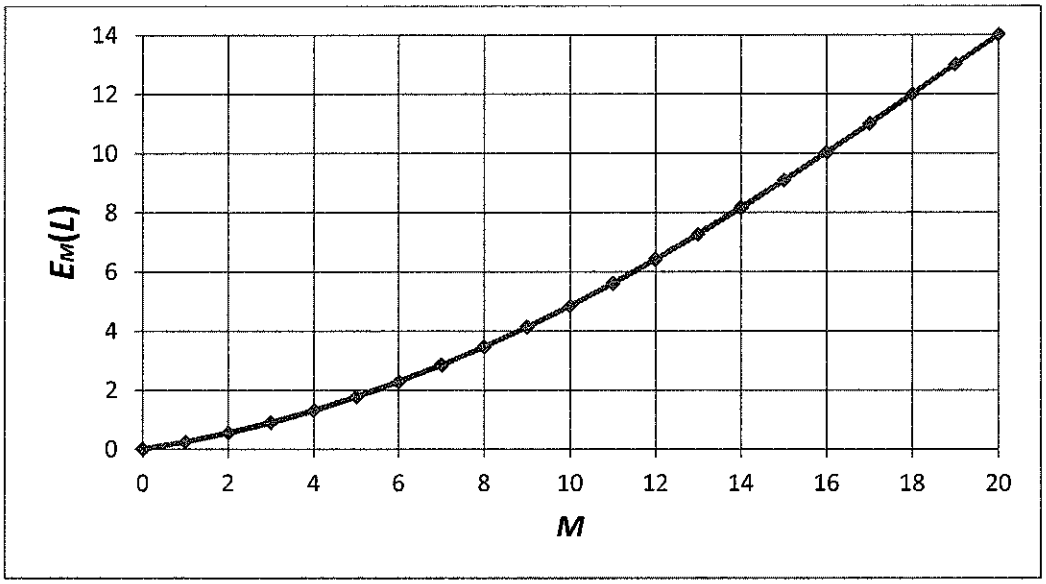

Figure 2. Expected Value of the Nonrandomized 90% Lower Bound $L(X)$.

If n is small compared to N, then the lower estimates for $\pi = M/N$ based on the Hypergeometric and the Binomial will be close. Table 2 contains results for several combinations of n, N and x.

Table 2. The Nonrandomized 90% Lower Bounds for $\pi = M/N$ when $n = 10$.

		x										
N		0	1	2	3	4	5	6	7	8	9	10
20	$L(x)$/N	0	0.050	0.100	0.150	0.250	0.350	0.450	0.550	0.650	0.750	0.850
50	$L(x)$/N	0	0.020	0.080	0.140	0.220	0.300	0.380	0.480	0.590	0.700	0.820
100	$L(x)$/N	0	0.020	0.060	0.130	0.200	0.280	0.370	0.460	0.570	0.680	0.810
500	$L(x)$/N	0	0.012	0.056	0.118	0.190	0.270	0.358	0.452	0.554	0.666	0.798
1000	$L(x)$/N	0	0.011	0.055	0.117	0.189	0.269	0.356	0.450	0.552	0.665	0.796
∞	$L_{binom}(x)$	0	0.010	0.055	0.116	0.188	0.267	0.354	0.448	0.550	0.663	0.794

3. Improvement through Randomization – an Example. The construction of an improved lower bound L_R for M through the introduction of artificial randomization is illustrated below. The construction is done in an iterative manner over steps indexed by $x = 0, 1, \ldots, n$; we use the name L_R for the result of each iteration. The final L_R is the same result that comes from inverting a family of *UMP* size α tests[3]. It has coverage probability $C_M(L_R) := P^*_{n,N,M}(L_R \le M)$ that satisfies

$$C_M(L_R) = 1 - \alpha \text{ for all } M = 0, 1, \ldots, N - 1. \tag{7}$$

For any lower estimate taking values in $\{0, 1, 2, \ldots, N\}$, $C_N(L_R) = 1$. The randomized version L_R that we construct produces lower bounds no less than those coming from the nonrandomized version L.

Below is a continuation of Example 1 to illustrate the construction of L_R. The constructions for $x = 0$ and 1 are given. The constructions for $x = 2, 3, 4$ and 5 follow the same pattern. Section 4 treats the general case.

One constructs a randomized version L_R through increasing the lower estimate by sacrificing the coverage probability that is in excess of 0.90 through artificial randomization. The construction begins at $x = 0$ and proceeds through $x = 1, \ldots, 5$. It may be useful to think of starting with L thought of as a randomized lower estimate L_R with degenerate randomization distributions. That is, at $x = 0$, the randomization distribution is degenerate at $L(0) = 0$; at $x = 1$, the randomization distribution is degenerate at $L(1) = 1$; at $x = 2$, the randomization distribution is degenerate at $L(2) = 3$; at $x = 3$, the randomization distribution is degenerate at $L(3) = 6$; at $x = 4$, the

[3] An alternative way to construct the randomized bounds is through the inversion of tests based on the convolution $Y = X + U$ where U is a *Uniform* (0, 1) random variable, independent of X. This idea is found in Lehmann (1959, p. 81). It is implemented in an R- program available from Professor Edwards.

randomization distribution is degenerate at $L(4) = 10$; at $x = 5$, the randomization distribution is degenerate at $L(5) = 14$.

Example 1 (cont). Construction of randomization distribution at $x = 0$. Here $L = 0$ with probability (*wp*) 1. Change the (conditional given $x = 0$) distribution to be $L = 0$ *wp* 0.9000 and $L = 1$ *wp* 0.1000. The change from a degenerate on the value 0 to this distribution has this effect: the new coverage probability function $C_M(L_R)$ takes the value 0.9000 at $M = 0$ and is seen to satisfy $C_M(L_R) = C_M(L) \geq 0.9000$, $M = 1, 2, \ldots, 20$. Note that $r_0(0) := 0.1000$ is the rejection probability on the boundary $x = 0$ of the *UMP* level $\alpha = 0.1000$ test of H_0: $M = 0$ v. H_1: $M > 0$.

Example 1 (cont). Construction of randomization distribution at $x = 1$. Since $C_1(L_R) = C_1(L) = F_1(1) = 1 > 0.9000$, we change the (conditional given $x = 1$) distribution on L from degenerate at $L = 1$ to 1 *wp* $(1 - r_1(1))$ and 2 *wp* $r_1(1)$ where $r_1(1)$ is chosen so that the new $C_1(L_R) = 0.9000$. Solving $F_1(0) + (1 - r_1(1))p_1(1) = 0.9000$ or the equivalent $\bar{F}_1(2) + r_1(1)p_1(1) = 0.1000$ leads to

$$r_1(1) = \frac{0.1000 - \bar{F}_1(2)}{p_1(1)} = 0.4000.$$

The change from a degenerate on the value 1 to the distribution 1 *wp* 0.6000 and 2 *wp* 0.4000 does not affect $C_0(L_R) = 0.9000$, produces $C_1(L_R) = 0.9000$, and retains $C_m(L_R) = C_m(L) \geq 0.9000$, $m = 2, 3, \ldots, 20$. Note that $r_1(1) = 0.4000$ is the rejection probability on the boundary $x = 1$ of the *UMP* level $\alpha = 0.1000$ test of H_0: $M = 1$ v. H_1: $M > 1$. The probability $r_1(1) = 0.4000$ on the lower bound $L = 2$ will now be split between the lower bounds $L = 2$ and $L = 3$. No split will change $C_m(L_R) = C_m(L)$, $m = 3, 4, \ldots, 20$;

yet one split will produce $C_2(L_R) = 0.9000$. The split is $(0.4000 - r_2(1))$ on $L = 2$ and $r_2(1)$ on $L = 3$ where $r_2(1)$ is the solution to

$F_2(0) + (0.6000 + 0.4000 - r_2(1))p_1(1) = 0.9000$ which is equivalent to $\bar{F}_2(2) + r_2(1)p_1(1) = 0.1000$. The solution is $r_2(1) = 0.1200$. Thus, the conditional distribution for the lower estimate given $x = 1$ is 1 *wp* 0.6000, 2 *wp* 0.2800 and 3 *wp* 0.1200 and it results in $C_0(L_R) = 0.9000$, $C_1(L_R) = 0.9000$, $C_2(L_R) = 0.9000$; $C_m(L_R) = C_m(L) \geq 0.9000$, $m = 3, 4, \ldots, 20$. Note that $r_2(1) = 0.1200$ is the rejection probability on the boundary $x = 1$ of the *UMP* level $\alpha = 0.1000$ test of H_0: $M = 2$ v. H_1: $M > 2$.

Table 3 gives the values determined for the randomizations in L_R at each x. The rows give the conditional on x randomization probabilities for choosing the value of the lower estimate. For example, if $x = 3$ is observed, the lower estimate is taken to be 6 *wp* 0.2668, 7 *wp* 0.3396, 8 *wp* 0.2165, 9 *wp* 0.1687 and 10 *wp* 0.0083. (Due to rounding, the conditional probabilities may not sum to 1 across a row.)

Table 3. The Randomizations for the 90% Lower Bound for M ($n = 5$, $N = 20$).

Observed x	L	L_R						
0	0	0 *wp* 0.9000	1 *wp* 0.1000					
1	1	1 *wp* 0.6000	2 *wp* 0.2800	3 *wp* 0.1200				
2	3	3 *wp* 0.3067	4 *wp* 0.3795	5 *wp* 0.2205	6 *wp* 0.0933			
3	6	6 *wp* 0.2668	7 *wp* 0.3396	8 *wp* 0.2165	9 *wp* 0.1687	10 *wp* 0.0083		
4	10	10 *wp* 0.3817	11 *wp* 0.2518	12 *wp* 0.1749	13 *wp* 0.1389	14 *wp* 0.0526		
5	14	14 *wp* 0.2256	15 *wp* 0.2581	16 *wp* 0.1613	17 *wp* 0.1044	18 *wp* 0.0696	19 *wp* 0.0476	20 *wp* 0.1333

Remark 2. The Geyer and Meeden (2005, (1.1b)) membership function values for their one-sided fuzzy interval estimate ϕ are the partial sums across the rows of Table 3. For $x = 2$, these are $\phi(2, 0.1000, M) = 0$ for $M = 0, 1, 2$; $\phi(2, 0.1000, 3) = 0.3067$; $\phi(2, 0.1000, 4) = 0.3067 + 0.3795 = 0.6862$; $\phi(2, 0.1000, 5) = 0.3067 + 0.3795 + 0.2205 = 0.9067$; $\phi(2, 0.1000, M) = 1$ for $M = 6, 7, \ldots, 20$.

Note that L_R differs from L by distributing an otherwise degenerate probability on $L(x)$ across the points $L(x)$, $L(x) + 1, \ldots, L(x + 1)$. Of course, $L_R \geq L$ *wp* 1. Figure 3 plots the expected values of L_R and L for Example 1.

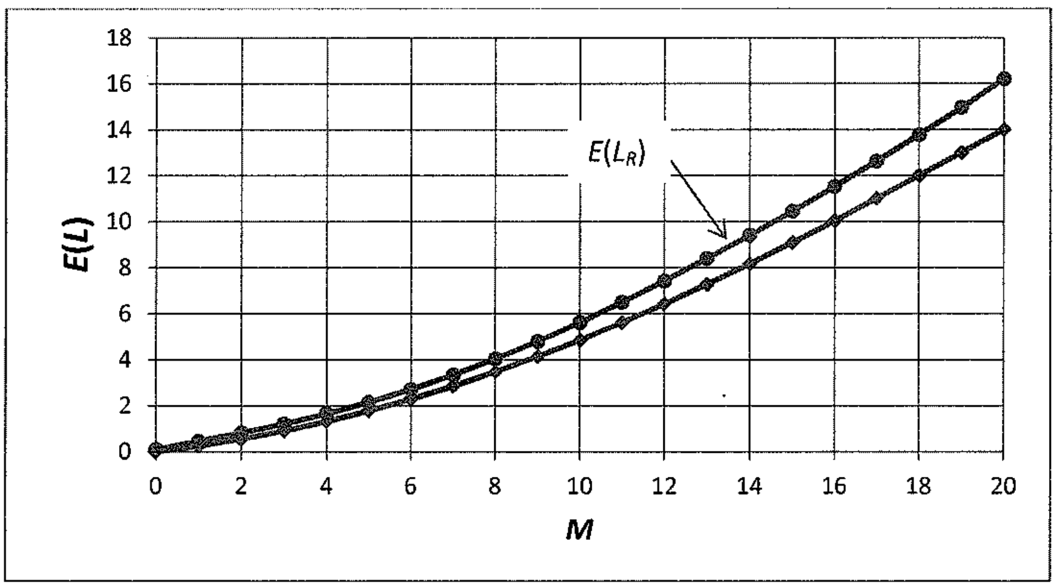

Figure 3. Expected Values of the Nonrandomized and Randomized 90% Lower Bounds.

4. Direct Construction of L_R in the General Case. To deploy L_R one takes the observation $X = x$ and computes the conditional probability distribution across lower estimates $L(x)$, $L(x) + 1$, …, $L(x + 1)$ coming from the inversion of the *UMP* tests. ($L(n) = N$ if and only if $n/N \geq 1 - \alpha$. In this case if $x = n$ is observed, no inversion is required.) Recall from Remark 1 that $L(x)$ is a strictly increasing function of x taking values in $\{0, 1, \ldots, N\}$. Letting $L(n + 1) := N$, we define the sets of consecutive integers

$$R(x) := \{L(x), L(x) + 1, \ldots, L(x + 1)\},\ x = 0, 1, \ldots, n. \tag{8}$$

If $L(n) = N$, the last set is the singleton set $\{N\}$. Note that

$$\begin{gathered} \alpha < \overline{F}_{L(x)}(x) \leq \overline{F}_{L(x)+1}(x) \leq \ldots \leq \overline{F}_{L(x+1)-1}(x) \\ \overline{F}_{L(x)}(x+1) \leq \overline{F}_{L(x)+1}(x+1) \leq \ldots \leq \overline{F}_{L(x+1)-1}(x+1) \leq \alpha . \end{gathered} \tag{9}$$

Thus, for each $m = L(x), L(x) + 1, \ldots, L(x + 1) - 1$, the *UMP* test of H_0: $M = m$ v. H_1: $M > m$ is

$$\phi_m(X) = [X \geq x+1] + r_m(x)[X = x] \tag{10}$$

where square brackets denote indicator functions and the boundary randomization rejection probability $r_m(x)$ is given by

$$r_m(x) := \frac{\alpha - \bar{F}_m(x+1)}{p_m(x)} \geq 0. \tag{11}$$

We point out that the sequence $r_m(x)$ is nonincreasing in m. Suppose that $r_{m+1}(x) > r_m(x)$. Then $\phi_{m+1} > \phi_m$ with positive probability under the Hypergeometric distribution P_{m+1} leads to the contradiction

$$\alpha = E_{m+1}\phi_{m+1} > E_{m+1}\phi_m \geq E_m\phi_m = \alpha \tag{12}$$

where E_{m+1} denotes expectation with respect to P_{m+1} and E_m denotes expectation with respect to P_m. In (12), the second inequality follows from the fact that ϕ_m is a nondecreasing function and Lehmann (1959, Lemma 2, p.74). Thus, $r_m(x) \geq r_{m+1}(x)$. Repeated application of the same argument shows that the rejection probabilities satisfy

$$r_{L(x)}(x) \geq r_{L(x)+1}(x) \geq ... \geq r_{L(x+1)-1}(x) \geq 0 . \quad (13)$$

(Thus, the Geyer and Meeden (2005, (1.1b)) membership function for the fuzzy confidence set is nondecreasing.) The randomization probabilities placed on $L(x)$, $L(x) + 1$, …, $L(x + 1)$ are shown in bottom row of Table 4.

Table 4. Randomizations for Lower Estimate L_R given $X = x$.

L	$L(x)$	$L(x) + 1$	….	$L(x + 1) - 1$	$L(x + 1)$
Conditional Reject Probs	$r_{L(x)}(x)$	$r_{L(x)+1}(x)$	….	$r_{L(x+1)-1}(x)$	
Conditional Prob Dist of L_R	$1 - r_{L(x)}(x)$	$r_{L(x)}(x) - r_{L(x)+1}(x)$	….	$r_{L(x+1)-2}(x) - r_{L(x+1)-1}(x)$	$r_{L(x+1)-1}(x)$

Example 2. Suppose that $\alpha = 0.1000$, $n = 20$, $N = 200$ and $x = 14$. Calculations yield $L(14) = 109$, $L(15) = 119$, and the randomization probabilities found in Table 5.

Table 5. Randomizations for 90% Lower Estimate given $X = 14$ ($n = 20$, $N = 200$).

L	109	110	111	112	113	114	115	116	117	118	119
Reject Probs*	.8747	.7559	.6450	.5427	.4454	.3545	.2693	.1881	.1104	.0363	
Conditional Prob Dist of L_R given $X = 14$	.1253	.1188	.1109	.1023	.0973	.0909	.0852	.0812	.0778	.0740	.0363

*Reminder: The *Reject Probs* are the boundary artificial randomization rejection probabilities of the *UMP* tests of 109 *v.* > 109, 110 *v.* > 110, etc.

The conditional expectation of L_R given $X = x$ is given by

$$E(L_R \mid X = x) = (1 - r_{L(x)}(x)) \cdot L(x) + \sum_{m=L(x)+1}^{L(x+1)-2} (r_m(x) - r_{m+1}(x)) \cdot L(m)$$
$$+ r_{L(x+1)-1}(x) \cdot L(x+1)$$
$$= L(x) + \sum_{m=L(x)}^{L(x+1)-1} r_m(x) \cdot \quad (14)$$

In case $x = n$ and $n/N \geq 1 - \alpha$, the last summation in (14) is from N to $N - 1$ and is taken as 0. Since $L(x) = E(L \mid X = x)$, the addition to conditional expectation due to the randomization is

$$\Delta E(L_R \mid X = x) = \sum_{m=L(x)}^{L(x+1)-1} r_m(x) . \quad (15)$$

The addition to the unconditional expectation of the lower bound coming from randomization is

$$\Delta E_M(L_R) = \sum_{x=0}^{n} p_M(x) \sum_{m=L(x)}^{L(x+1)-1} r_m(x), \; M = 0, 1, 2, \ldots, N. \quad (16)$$

Example 3. See Example 2 of Section 4. The addition (15) to the conditional expectation of the lower estimate given $X = 14$ can be calculated from Table 5 by simply adding across its second row. The result is $\Delta E(L_R \mid X = 14) = 4.2$.

Remark 3. Conditional on $X = x$ with $x = 1, 2, \ldots, n - 1$, the randomization distribution on $\{L(x)/N, (L(x) + 1)/N, \ldots, L(x + 1)/N\}$ converges to a probability distribution on the interval $(L_{binom}(x), L_{binom}(x+1))$ as $N \to \infty$. The limit cumulative distribution function is

$$F(\pi \mid X = x) = 1 - \frac{\alpha - \overline{B}_\pi(x+1)}{b_\pi(x)} = 1 - \frac{\alpha - 1 + B_\pi(x)}{b_\pi(x)}, \; L_{binom}(x) \leq \pi \leq L_{binom}(x+1) \quad (17)$$

where b_π is the probability mass function, $\overline{B}_\pi$ is the right-tail probability, and B_π is the left tail probability for $B(n, \pi)$. If $X = 0$, the limit distribution places mass $1 - \alpha$ on 0. If $X = n$, the limit distribution places mass α on 1. The other distributions are continuous on their intervals of support.

Proof. Fix π in the interval $(L_{binom}(x), L_{binom}(x+1))$ and let $L(x), L(x) + 1, \ldots, L(x) + d_N = L(x + 1)$ be the support of the randomized lower estimate L_R based on the Hypergeometric and given $X = x$. Let F_N denote cumulative distribution of L_R/N. Take $k_N \leq d_N$ to be the integer for which

$$\max\{L_{binom}(x), (L(x)+k_N)/N\} \leq \pi \leq \min\{L_{binom}(x+1), (L(x)+k_N+1)/N\}, \quad N = 1, 2, \ldots \qquad (18)$$

The endpoints of (18) converge to π as $N \to \infty$. Examination of the cumulative distribution function F_N evaluated at $(L(x) + k_N)/N$ and at $(L(x) + k_N + 1)/N$ and noting the collapsing sum of the probabilities in the bottom row of Table 4 shows that $F_N(\pi) \to F(\pi)$ as $N \to \infty$ follows from the convergence of the Hypergeometric to the Binomial. The limit distribution (17) is the distribution for the randomized lower estimate for the Binomial π developed from the convolution approach of Lehmann (1959, Example 7, p. 81). □

Figure 4 is a graph of the limit distributions (17) in the case $n = 5$. The endpoints of the intervals are found in the bottom row of Table 1.

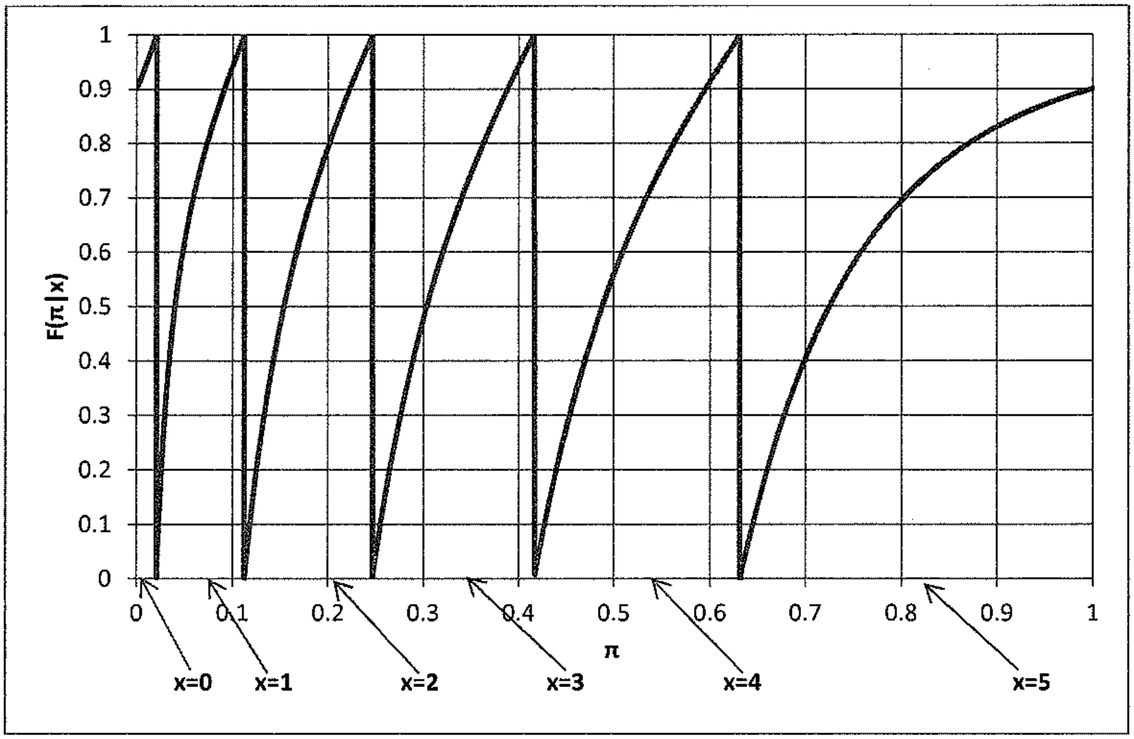

Figure 4. Distributions for Randomized Lower Estimates for π of $B(5, \pi)$.

5. Applications and Final Remarks. Guidelines from the Center for Medicare & Medicaid Services (2003, p. 6) include "In most situations, the lower limit of a one-sided 90 percent confidence interval should be used as the amount of overpayment to be demanded for recovery from the physician or supplier." The Edwards, et al (2005) *minimum sum method* (*MSM*) uses the nonrandomized lower bound L discussed in Section 2 and is very effective when all population payments are about equal, or when the denial rate is very high. A way to adapt *MSM* to cases where payments vary is presented in Gilliland and Feng (2010) and Edwards et al (2010). The applications below illustrate the extent to which randomized lower bound can increase the recovery demand by the insurer.

In these applications, we suggest a modification of the randomized lower bound L_R based on the reality of the sometimes contentious overpayment recovery efforts by an insurer. We suggest that if $x = 0$, take the lower bound for M to be 0. With this modification of L_R the coverage probability is increased to 1 if $M = 0$. In addition, the situation of attempting a small but positive recovery (10% of the minimum payment) when no overpayments are observed in the sample is avoided.

Application 1. A population of $N = 20$ payments made by an insurer to a medical device supplier are each for $4,000 and are either valid or not[4]. A simple random sample of $n = 5$ payments are investigated and in $x = 3$ cases the claims are found to be invalid. Suppose that the insurer seeks recovery of the overpayment at the 90% confidence lower estimate. The nonrandomized lower estimate of the number M of claims in the population that are invalid is $L(3) = 6$. Thus, use of L will result in seeking recovery of $24,000. On the other hand, Table 4 shows that conditional on $X = 3$, the expected value of the randomized lower estimate is 0.2668(6) + 0.3396(7) + 0.2165(8) + 0.1688(9) + 0.0083(10) = 7.31. Thus, use of L_R will result in seeking recovery of an expected $29,240. □

Application 2. A population of $N = 200$ payments made by an insurer to a medical device supplier are each for $4,000 and are either valid or not[5]. A simple random sample of $n = 20$ payments are investigated, and in $x = 14$ cases the claims are found to be invalid. Suppose that the insurer seeks recovery of the overpayment at the 90% confidence lower estimate. The nonrandomized lower estimate of the number M of

---

[4] According to the AARP Bulletin (November 2009, p. 3) the average cost to Medicare for a motorized wheelchair in 2007 was $4,018.

[5] Populations of about this size are not rare as targets for auditing. Edwards et al (2003) gives four payment populations ranging in size from 126 to 292.

claims in the population that are invalid is $L(14) = 109$. Thus, use of L will result in seeking recovery of $436,000. From Table 4 the conditional expected value of L_R given $X = 14$ is determined to be 113.2. Thus, use of L_R will result in seeking recovery of an expected $452,800. □

Application 3. Actual Medicare Audit. This population is Test Population 1 in Edwards et al (2005). An actual audit of a simple random sample of $n = 30$ payments from the population of $N = 292$ payments showed all sample payments completely in error. The 90% nonrandomized lower bound for the population number of payments M that are totally in error is $L(30) = 272$. By (15), $\Delta E(L_R|X = 30) = 7.9$. In this case, the addition of 7.9 to the lower bound for M adds 7.9 x $4,042 = $31,932 to the recoupment demand. □

The payment population of Application 3 is depicted in Figure 5a below. The audit of the simple random sample of $n = 30$ payments spurred the development of the minimum sum method because the nominal 90% lower confidence bound based on the CLT (using the simple expansion estimator) produced an illegal lower confidence bound for "total overpayment": one greater than the total payment in the population. In practice, this problem has often been dealt with by truncating the lower bound's value down to the total population payment amount. This truncation does not, however, repair the fundamental problem of the CLT methods, that their achieved confidence level for high error rates in populations like these can drop far below the nominal level because of left-skew in the overpayment population. Figure 5b shows the Monte-Carlo estimated confidence level for this example for the two most widely used CLT methods, the simple expansion estimator and the ratio estimator, as well as the ordinary and randomized minimum sum methods, plotted against the actual population error rate. We see that the CLT methods are not valid when the payment error rate exceeds 50-60%; the ordinary minimum sum method is conservative under all error rates, while the randomized minimum sum method gives essentially exactly a 90% lower bound for all error rates on (0,1). Figure 5c shows the average overpayment recovery of these methods plotted only in those regions where the method provides the nominal confidence level. We see that the minimum sum methods' expected recovery exceeds or matches that of the CLT methods wherever the latter are valid; also that the randomized minimum sum method gives a consistently higher average recovery compared to the simple minimum sum method.

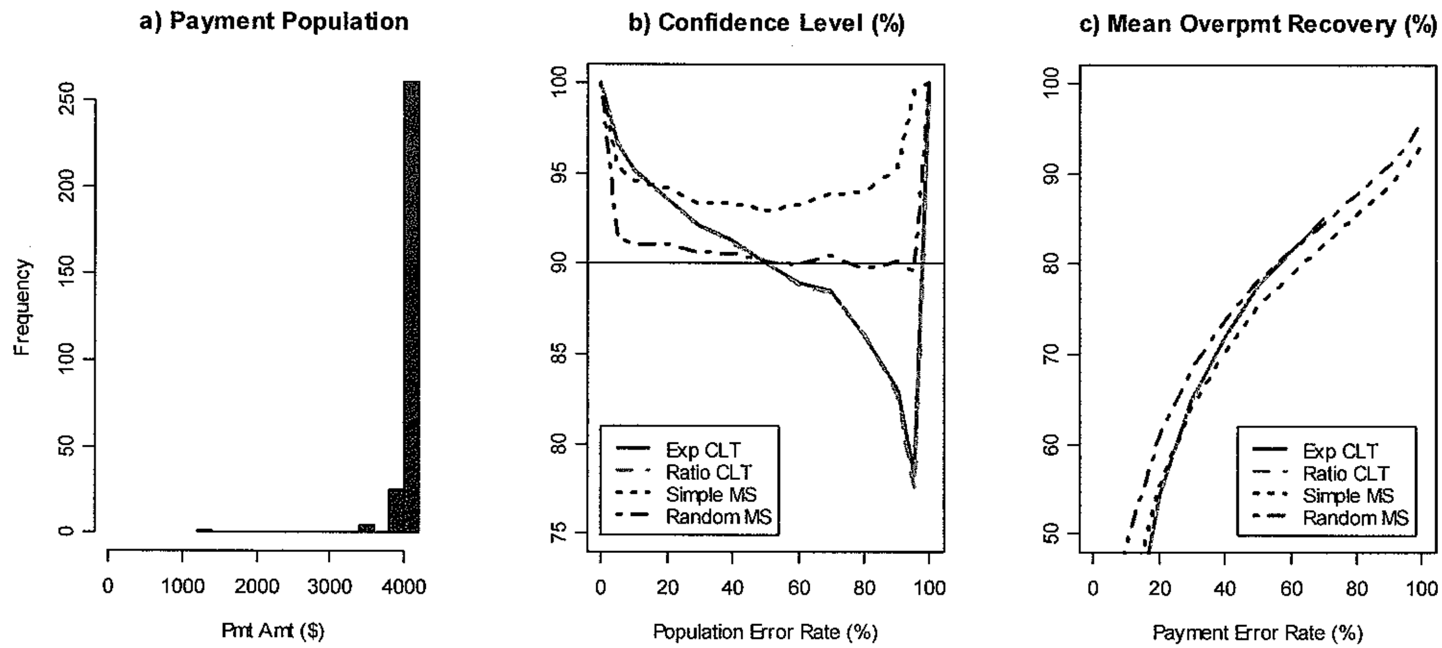

Figure 5: (a) histogram of population payment amounts, N = 292 motorized wheelchairs; (b) Monte Carlo estimated confidence levels for nominal 90% lower confidence bounds for the CLT expansion estimator, the CLT ratio estimator, the simple minimum sum method, and the randomized minimum sum method; (c) average overpayment recovery for the same four methods over their valid range. Simple random samples of size n=30. Monte Carlo estimates have standard errors < ½ %.

In using L_R, it seems reasonable that the medical service provider being audited have a choice between paying the conditional expected value of L_R and casting its lot with a supervised and agreed upon method of executing the randomization.

In a long term relationship between health care provider and insurer, recovery of overpayments at the 90% confidence lower estimate by itself provides no incentive to the health care provider to avoid overpayments. For the sake of discussion, suppose that a valid, risk symmetric, 80% confidence interval estimate of total overpayment is $500,000 ± $100,000, where $500,000 is an unbiased estimate. Then $400,000 is a 90% confidence lower estimate of total overpayment. If this is the recovery figure, then across independent repetitions of these audits, taken for simplicity to produce the same interval estimate each time, the health care provider will pay back only 80 cents on the dollar for overpayments by the insurer. With a valid study, recovery at a lower estimate with high confidence level does not disadvantage the health care provider, no matter how large the margin of error. On the other hand, recovery at the unbiased estimate $500,000 will be fair in the long run.

This argues for the use of a point estimate in setting a recovery figure. Naturally, even an unbiased point estimate is challenged when its margin of error is large because of the large risk born by the provider and because the insurer designed the audit and is taken as responsible for the large margin of error. If the margin of error is very large compared to the point estimate, there may be reluctance on the part of the court or administrative law judge to order recovery at the point estimate. On the other hand, the health care provider can hardly complain about the use of a lower estimate where it benefits from a large margin of error.

REFERENCES

The Case of the Expensive Wheelchair, *AARP Bulletin*, p. 3, November 2009.

Brown, L. D., Cai, T. T., and DasGupta, A., Interval Estimation for a Binomial Proportion, *Statistical Science*, **16**, pp. 101-133, 2001.

Brown, L. D., Cai, T. T., and DasGupta, A., Confidence Intervals for a Binomial Proportion and Asymptotic Expansions, *Ann. Statist.* **30**, pp. 160-201, 2002.

Brown, L. D., Cai, T. T., and DasGupta, A., Comment: Fuzzy and Randomized Confidence Intervals and *P*-Values, *Statistical Science*, **20**, pp. 375-379, 2005.

Center for Medicare and Medicaid Services, Program Memorandum Carriers Change Request 1363, CMS-Pub. 60B, U.S. Department of Health and Human Services, March 21, 2003.

Edwards, D., Ward-Besser, G., Lasecki, J., Parker, B., Wieduwilt, K., Wu, F. and Moorhead, P., The Minimum Sum Method: a Distribution-free Sampling Procedure for Medicare Fraud Investigations, *Health Services & Outcomes Research Methodology*, **4**, 241-263, 2005.

Edwards, D., Gilliland, D., Ward-Besser, G. and Lasecki, J., A Primer on Pennysampling, Preprint (2010).

Geyer, C. J. and Meeden, G. D., Fuzzy and Randomized Confidence Intervals and *P*-Values, *Statistical Science*, **20**, pp. 365-374, 2005.

Gilliland, D. and Feng, W., An Adaptation of the Minimum Sum Method, *Health Services & Outcomes Research Methodology*, to appear, 2010.

Lehmann, E. L., *Testing Statistical Hypotheses*, Wiley, New York, 1959.

Lehmann, E. L. and Romano, J. P., *Testing Statistical Hypotheses, 3rd Edition*, Wiley, New York, 2005.

Stapleton, J. H., *Linear Statistical Models, 2nd Edition*, Wiley, New York, 2009.