ANALYTICAL PROPERTIES OF BAYESIAN COX-SNELL BOUNDS IN AUDITING

by J.J.A. Moors and M.J.B.T. Janssens

No. 87.01

Contributed paper
Third Valencia International Meeting on Bayesian Statistics 1987
ANALYTICAL PROPERTIES OF BAYESIAN COX-SNELL BOUNDS IN AUDITING

by J.J.A. Moors and M.J.B.T. Janssens

Summary. A class of Bayesian models for auditing problems was proposed by COX & SNELL (1979). Applied to a given population each model gives a posterior upper bound for the relative overstatement error. GODFREY & NETER (1984) and NETER & GODFREY (1985) thoroughly investigated these upper bounds. They applied the bounds to a number of theoretical populations and studied their behaviour by means of simulation. It is shown here how the analytical properties of the Cox-Snell models can be further exploited to derive the exact distributions of the upper bounds.

1. Introduction

Consider a population of recorded values in which errors may occur, but only overstatement errors. An auditor wants to find an upper bound for the relative overstatement error \( \psi \), that is the total of the errors expressed as a fraction of the total of the recorded values. From past experience the auditor usually has fairly good insight in the characteristics of the population at hand. Therefore, it seems natural to make use of this prior knowledge and to consider this problem from a Bayesian point of view.

Up to now the most flexible class of Bayesian models was proposed by COX & SNELL (1979). First, they split \( \psi \) into two separate factors: \( \psi = \varphi/\lambda \), where the parameter \( \varphi \) denotes the fraction of elements in error and \( 1/\lambda \) is the mean relative overstatement error per recorded value. From the Bayesian viewpoint \( \varphi \) and \( \lambda \) are seen as random variables; their (prior) distributions reflect the auditor's prior knowledge. In the Cox-Snell models \( \varphi \) and \( \lambda \) are independent, both having a prior distribution from the class of (two-parameter) gamma distributions. Denoting the gamma distribution with density \( b^{b-1} x^{b-1} \exp(-ax)/\Gamma(b) \), \( x > 0 \), by \( \Gamma(a,b) \), an obvious notation gives

\[
(1.1) \quad (\varphi, \lambda) \sim \Gamma(a,b) \times \Gamma(c,d)
\]
So, in any particular case the auditor has to specify (positive) values for the four parameters $a$, $b$, $c$ and $d$, indeed giving much flexibility. Standard statistical theory shows that the ratio $(a/b)\varphi/[c/d]\lambda$ has the well-known F-distribution with the pair $(2b,2d)$ as degrees of freedom:

\begin{equation}
\frac{ad}{bc} \varphi \sim F_{2b}^{2d}
\end{equation}

The auditor's prior knowledge, formalized by (1.2), can be combined with the empirical information obtained from a sample. Assume that the population is sampled with probability proportional to recorded value (monetary unit sampling). Let $m$ denote the number of (overstatement) errors found in a sample of size $n$ and let $z$ denote the total overstatement error in the sample. Then the posterior distribution of $\varphi$ and $\lambda$ - given these sample results - can be shown to consist of two independent gamma distributions again:

\begin{equation}
(\varphi, \lambda) | (m,z) \sim \Gamma(a+n,b+m) \times \Gamma(c+z,d+m)
\end{equation}

Once again, a posterior F-distribution results for the random variable $\varphi$:

\begin{equation}
\frac{(a+n)(d+m)}{(b+m)(c+z)} \varphi | (m,z) \sim F_{2d}^{2(b+m)}
\end{equation}

A posterior 95% confidence upper bound for the true value of $\psi$ follows immediately:

\begin{equation}
\lambda = \frac{(b+m)(c+z)}{(a+n)(d+m)} F_{2(b+m)}^{2(d+m)}; 0.95
\end{equation}

where the last factor denotes the 95th percentile of the F-distribution. Note that the upper bound $\lambda$ indeed incorporates both the sample data $n$, $m$ and $z$ and the auditor's prior knowledge ($a$, $b$, $c$ and $d$).

These results were essentially derived by COX & SNELL (1979), although their final formula contained an error. Since their derivation was rather sketchy besides, MOORS (1983) presented a corrected, more lucid derivation. Further, this last paper described an analytical method to find the properties of $\lambda$ when applied to a theoretical distribution. Since
this is the subject of the present paper as well, we will return to it in due course.

If an appropriate Cox-Snell model has been specified (that is, if the constants $a$, $b$, $c$ and $d$ have been chosen) the behaviour of the upper bound $\lambda$ - when applied to a given population - can be studied by means of simulation. If a sample has been drawn from the population, the value of $\lambda$ can be calculated. After repeating this a large number of times, the percentage of cases can be found in which $\lambda$ exceeds the true value of $\psi$ in the population. This percentage is called the coverage; ideally, it is at least 95%, since $\lambda$ is a 95%-confidence bound for $\psi$. However, if the population contains more and/or larger errors than according to the prior model, the coverage may be (much) lower than 95%. If the coverage is nearly 95%, even for populations that deviate greatly from the prior model, the prior model is called robust.

GODFREY & NETER (1984) and NETER & GODFREY (1985) extensively studied the robustness of the Cox-Snell models. Their main conclusion was that some models indeed have very good robustness properties. To indicate the size of their calculations, the 1984 paper considered ten Cox-Snell models and each of them was applied to 21 theoretical populations. Since the number of replications was 500 and the sample size 100 throughout, the total number of simulated observations equals $10 \times 21 \times 500 \times 100$, or over $10^7$. The 1985 paper presents some more.

The present paper offers a possibility to avoid simulation altogether by presenting a more analytical derivation of the distribution of the upper bound $\lambda$, when applied to a given theoretical population. Among other things, the exact value of the coverage - of which simulation can only give an estimate - follows at once. Some of the results presented here can be found (in Dutch) in MOORS (1986).

The organisation of the paper is as follows. Section 2 outlines the general method and discusses the choice of both the prior models and the populations to which it was applied. In Section 3 we give a detailed description of the calculations and present the results. Since only populations were selected in Section 2 for which the analysis is relatively simple, we indicate in Section 4 how the more difficult cases can be treated. The final Section 5 briefly discusses the results.
2. General method and model/population selection

For a given population, m and z are determined by the sampling results and can be viewed therefore as the outcomes of two random variables M and Z, respectively. For a given prior model, the upper bound \( \lambda \) in (1.5) only depends on m and z. Before sampling and observing have been carried out, the upper bound therefore is a random variable as well, to be denoted by L. It may be written as

\[
L = \frac{(b+M)(c+Z)}{(a+n)(d+M)} F^2(b+M) \; 2(d+M); 0.95
\]

The approach of this paper is based on the observation that, for given \( M = m \), L is a linear transformation of Z. Hence the conditional distribution of \( L \) - given \( M = m \) - easily follows from the conditional distribution of Z (given \( M = m \)). By taking into account the distribution of M, the unconditional distribution of L can be derived.

If the distribution function of L is denoted by H, we find, more explicitly, by conditioning with respect to M:

\[
H(\lambda) = P(L \leq \lambda) = \sum_{m=0}^{n} P(L \leq \lambda | M = m)P(M = m)
\]

This can be rewritten by means of (2.1) as

\[
H(\lambda) = \sum_{m=0}^{n} P(Z \leq \lambda - c | M = m)P(M = m)
\]

where \( e_m \) is defined as

\[
e_m = \frac{(a+n)(d+m)}{b+m} F^2(b+m) \; 2(d+m); 0.95
\]

Denoting the conditional distribution function of Z, given \( M = m \), by \( G_m \), so that \( G_m(z) = P(Z \leq z | M = m) \), (2.2) can be simplified to

\[
H(\lambda) = \sum_{m=0}^{n} G_m(e_m \lambda - c)P(M = m)
\]
Now, for a fully known population $G_m$ can be calculated, while $P(M = m)$ follows from the sampling plan; hence, the distribution function $H$ of $L$ can be found for any given population - at least in principle.

More specifically, the expectation $E(L)$ and variance $V(L)$ can be derived. The general formula

\[(2.5) \quad E(L) = E[E(L|M)]\]

leads to

\[(2.6) \quad E(L) = \sum_{m=0}^{n} P(M = m)E(L|M = m) = \sum_{m=0}^{n} P(M = m)[c + E(Z|M = m)]/c_m\]

while the variance follows from the general formula

\[(2.7) \quad V(L) = V[E(L|M)] + E[V(L|M)]\]

In the following, populations will be considered for which the error distribution is fully known. Let one recorded value be drawn with probability proportional to size; its relative overstatement will be denoted by $S$. Then the probability distribution of $S$ is known as well; in the sequel this distribution will be given in stead of the underlying error distribution in the population. Since sampling with replacement will be assumed throughout, the distribution of $Z = \sum S_i$ can be found by the standard technique of calculating convolutions. Throughout the paper $S$ will be assumed to have a point mass at zero of size $1-\varphi$; at the moment, the conditional distribution of $S$, given $S > 0$ will be assumed to have a density, to be denoted by $g$. (This last assumption will be dropped in Section 4, where an additional point mass will be allowed.) Summarizing, $S$ will be assumed to have the mixed distribution:

\[(2.8) \quad \begin{cases} P(S = 0) = 1 - \varphi \\ P(s < S < s + ds) = \varphi g(s)ds, \; s > 0 \end{cases}\]

where $ds$ denotes an infinitesimal length.
In their simulation study, GODFREY & NETER (1984) considered 21 theoretical populations, all of them featuring a gamma distribution - more specifically, \( x^2 \) - and exponential distributions. In fact, however, their populations 4-6 in Table 5 are identical to the populations 10-12, respectively. Indeed 4-6 use the exponential distribution \( Ne(5) = \Gamma(5,1) \), while 10-12 make use of the distribution \( x^2_2 = \Gamma(1,1) = Ne(1) \) after division by 10; the identity follows from the general property

\[
X \sim Ne(\lambda) \Rightarrow \alpha X \sim Ne(\lambda/\alpha), \quad \alpha > 0
\]

Note that \( P(S > 1) \) is positive for all gamma distributions, although \( S \) can not exceed 1 by definition. Therefore all densities were truncated at 1.

From these 21 (or 18) populations we chose the populations 1 and 4 in our study, they have the advantage that \( P(S > 1) \) is small: \( 10^{-9} \) and 0.0067, respectively. Hence, truncation at 1 is not necessary, which greatly facilitates the analysis. Section 4 briefly considers the difficulties arising from such a truncation. Table 1 summarizes some properties of these two populations. This table further includes as population 22 the case that \( g \) has the uniform distribution on \([0,1]\), considered by MOORS (1983).

**Table 1 Populations under study**

<table>
<thead>
<tr>
<th>Number</th>
<th>( \varphi )</th>
<th>( g )</th>
<th>( \psi ) (in %)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.01</td>
<td>Ne(20)</td>
<td>0.0495</td>
</tr>
<tr>
<td>4</td>
<td>0.01</td>
<td>Ne(5)</td>
<td>0.198</td>
</tr>
<tr>
<td>22</td>
<td>0.01</td>
<td>U(0,1)</td>
<td>0.495</td>
</tr>
</tbody>
</table>

Note that for population 4 the value of \( \psi \) slightly differs from Godfrey and Neter's value 0.193, which is caused by the truncation at 1. Further,
the range of \( y \)-values is rather wide, representing populations with quite different error distributions.

Now we turn to the choice of the Cox-Snell models to be studied here. GODFREY & NETER (1984) considered ten of these, defined by a specific choice of the parameters in (1.1). We picked the model CS1, CS4 and CS10, representing quite dissimilar prior beliefs. These three are summarized in Table 2. Note that CS10 is precisely the 'pessimistic' case treated in MOORS (1983).

### Table 2. Cox-Snell models under study

<table>
<thead>
<tr>
<th>Number</th>
<th>a</th>
<th>b</th>
<th>( E(\phi) )</th>
<th>( \sigma_\phi )</th>
<th>c</th>
<th>d</th>
<th>( E(1/\lambda) )</th>
<th>( \sigma_{1/\lambda} )</th>
<th>( E(\psi) )</th>
<th>( \sigma_\psi ) (in %)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CS1</td>
<td>400</td>
<td>4</td>
<td>0.01</td>
<td>0.005</td>
<td>0.25</td>
<td>6</td>
<td>0.05</td>
<td>0.025</td>
<td>0.05</td>
<td>0.0375</td>
</tr>
<tr>
<td>CS4</td>
<td>400</td>
<td>4</td>
<td>0.01</td>
<td>0.005</td>
<td>3.4</td>
<td>18</td>
<td>0.2</td>
<td>0.05</td>
<td>0.2</td>
<td>0.1146</td>
</tr>
<tr>
<td>CS10</td>
<td>10</td>
<td>1</td>
<td>0.1</td>
<td>0.1</td>
<td>2</td>
<td>6</td>
<td>0.4</td>
<td>0.2</td>
<td>4</td>
<td>4.8990</td>
</tr>
</tbody>
</table>

For each of these models the distribution of the corresponding (random) bound \( L \) in (2.1) will be derived for any of the populations in Table 1. So in all, nine probability distributions will be presented in the next section.

### 3. Results

All calculations were done on an IBM PC. First of all, 95-percentiles of F-distributions were calculated in four decimals, to be used in (2.3). The sample size \( n = 100 \) was used throughout.

For populations 1 and 4 the density \( g \) corresponds to distribution \( Ne(\lambda) = \Gamma(\lambda,1) \), so that its convolution is \( \Gamma(\lambda,2) \). Hence for \( m > 0 \), the distribution function \( G_m \) corresponds to \( \Gamma(\lambda,m) \); of course for \( m = 0 \), \( Z = 0 \) with probability 1, so that \( L \) is degenerated as well. By consequence
\( G_m(z) \) and hence \( H(\ell) \) in (2.4) can be calculated for each of the three prior models in Table 2. Table 3 presents these conditional and absolute distribution functions for one case. Note that \( P(M \geq 7) = 0.0001 \), so that the distributions for \( M \geq 7 \) are irrelevant.

**Table 3. Distribution functions \( G_m \) and \( H \) for population 1 and model CS1**

<table>
<thead>
<tr>
<th>( \ell )</th>
<th>( G_1(\ell) )</th>
<th>( G_2(\ell) )</th>
<th>( G_3(\ell) )</th>
<th>( G_4(\ell) )</th>
<th>( G_5(\ell) )</th>
<th>( G_6(\ell) )</th>
<th>( H(\ell) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.093</td>
<td>0.0034</td>
<td>0.0060</td>
<td>0.0015</td>
<td>0.0004</td>
<td>0.0001</td>
<td>0.0000</td>
<td>0.0025</td>
</tr>
<tr>
<td>0.103</td>
<td>0.1331</td>
<td>0.1434</td>
<td>0.0450</td>
<td>0.0132</td>
<td>0.0037</td>
<td>0.0010</td>
<td>0.5501</td>
</tr>
<tr>
<td>0.113</td>
<td>0.6602</td>
<td>0.3244</td>
<td>0.1532</td>
<td>0.0616</td>
<td>0.0228</td>
<td>0.0078</td>
<td>0.6838</td>
</tr>
<tr>
<td>0.123</td>
<td>0.8016</td>
<td>0.5263</td>
<td>0.2978</td>
<td>0.1501</td>
<td>0.0688</td>
<td>0.0291</td>
<td>0.7804</td>
</tr>
<tr>
<td>0.133</td>
<td>0.8842</td>
<td>0.6723</td>
<td>0.4478</td>
<td>0.2666</td>
<td>0.1443</td>
<td>0.0718</td>
<td>0.8490</td>
</tr>
<tr>
<td>0.153</td>
<td>0.9605</td>
<td>0.8547</td>
<td>0.6955</td>
<td>0.5176</td>
<td>0.3542</td>
<td>0.2242</td>
<td>0.9305</td>
</tr>
<tr>
<td>0.183</td>
<td>0.9921</td>
<td>0.9616</td>
<td>0.8966</td>
<td>0.7953</td>
<td>0.6755</td>
<td>0.5275</td>
<td>0.9794</td>
</tr>
<tr>
<td>0.213</td>
<td>0.9984</td>
<td>0.9906</td>
<td>0.9694</td>
<td>0.9277</td>
<td>0.8611</td>
<td>0.7706</td>
<td>0.9942</td>
</tr>
<tr>
<td>0.243</td>
<td>0.9997</td>
<td>0.9978</td>
<td>0.9917</td>
<td>0.9774</td>
<td>0.9501</td>
<td>0.9059</td>
<td>0.9984</td>
</tr>
<tr>
<td>0.273</td>
<td>0.9999</td>
<td>0.9995</td>
<td>0.9979</td>
<td>0.9935</td>
<td>0.9839</td>
<td>0.9660</td>
<td>0.9996</td>
</tr>
<tr>
<td>0.303</td>
<td>1.0000</td>
<td>0.9999</td>
<td>0.9995</td>
<td>0.9982</td>
<td>0.9952</td>
<td>0.9888</td>
<td>0.9999</td>
</tr>
<tr>
<td>0.333</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.9999</td>
<td>0.9996</td>
<td>0.9987</td>
<td>0.9966</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

* point mass 0.3661 at 0.0950

For population 22 density \( g \) corresponds to the uniform distribution on \([0,1]\). By taking convolutions \( G_m \) can be found:

\[
G_2(z) = \begin{cases} 
  z^2/2, & 0 \leq z \leq 1 \\
  (-z^2+4z-2)/2, & 1 \leq z \leq 2 
\end{cases}
\]

\[
G_3(z) = \begin{cases} 
  z^3/6, & 0 \leq z \leq 1 \\
  (-2z^3 + 9z^2 - 9z + 3)/6, & 1 \leq z \leq 2 \\
  (z^3 - 9z^2 + 27z - 21)/6, & 2 \leq z \leq 3 
\end{cases}
\]
$G_4(z) = \begin{cases} 
\frac{z^4}{24}, & 0 \leq z \leq 1 \\
\frac{(-3z^4 + 16z^3 - 24z^2 + 16z - 4)}{24}, & 1 \leq z \leq 2 \\
\frac{(3z^4 - 32z^3 + 120z^2 - 176z + 92)}{24}, & 2 \leq z \leq 3 \\
\frac{(-z^4 + 16z^3 - 96z^2 + 256z - 232)}{24}, & 3 \leq z \leq 4 
\end{cases}$

For $m = 5$ and $m = 6$ the normal approximation $N(m/2, m/12)$ has been used; cf. MOORS (1983). Table 4 gives the detailed results for model CS1.

Table 4. Distribution functions $G_m$ and $H$ for population 22 and model CS1.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0.093</td>
<td>0.0002</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.0001*</td>
</tr>
<tr>
<td>0.133</td>
<td>0.1078</td>
<td>0.0067</td>
<td>0.0003</td>
<td>0</td>
<td>0.0001</td>
<td>0</td>
<td>0.4072</td>
</tr>
<tr>
<td>0.173</td>
<td>0.2154</td>
<td>0.0255</td>
<td>0.0022</td>
<td>0.0002</td>
<td>0.0003</td>
<td>0.0001</td>
<td>0.4505</td>
</tr>
<tr>
<td>0.213</td>
<td>0.3230</td>
<td>0.0563</td>
<td>0.0070</td>
<td>0.0007</td>
<td>0.0005</td>
<td>0.0001</td>
<td>0.4963</td>
</tr>
<tr>
<td>0.313</td>
<td>0.5920</td>
<td>0.1864</td>
<td>0.0414</td>
<td>0.0072</td>
<td>0.0022</td>
<td>0.0005</td>
<td>0.6220</td>
</tr>
<tr>
<td>0.413</td>
<td>0.8610</td>
<td>0.3921</td>
<td>0.1253</td>
<td>0.0314</td>
<td>0.0083</td>
<td>0.0021</td>
<td>0.7650</td>
</tr>
<tr>
<td>0.513</td>
<td>1</td>
<td>0.6476</td>
<td>0.2775</td>
<td>0.0913</td>
<td>0.0259</td>
<td>0.0072</td>
<td>0.8739</td>
</tr>
<tr>
<td>0.613</td>
<td>1</td>
<td>0.8406</td>
<td>0.4781</td>
<td>0.2026</td>
<td>0.0676</td>
<td>0.0212</td>
<td>0.9236</td>
</tr>
<tr>
<td>0.813</td>
<td>1</td>
<td>0.9999</td>
<td>0.8490</td>
<td>0.5509</td>
<td>0.2771</td>
<td>0.1166</td>
<td>0.9816</td>
</tr>
<tr>
<td>1.013</td>
<td>1</td>
<td>1</td>
<td>0.9888</td>
<td>0.8637</td>
<td>0.6221</td>
<td>0.3613</td>
<td>0.9959</td>
</tr>
<tr>
<td>1.213</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.9847</td>
<td>0.8875</td>
<td>0.6852</td>
<td>0.9993</td>
</tr>
<tr>
<td>1.413</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.9999</td>
<td>0.9828</td>
<td>0.9065</td>
<td>0.9999</td>
</tr>
</tbody>
</table>

* point mass 0.3661 at 0.0950

The distribution functions $H$ for the nine combinations are shown in Figures 1-3; detailed tables are available on request. Some parameters of these distributions will be given below. Table 5 presents expectation and standard deviation of $L$, found by means of (2.6) and (2.7).
Figure 1. Distribution functions of $L$ for Cox-Snell model CS1.
Figure 2. Distribution functions of L for Cox-Snell model CS4.
Figure 3. Distribution functions of L for Cox-Snell model CS10.
Table 5. Expectation and standard deviation of L (in %).

<table>
<thead>
<tr>
<th>model</th>
<th>CS1</th>
<th>CS4</th>
<th>CS10</th>
</tr>
</thead>
<tbody>
<tr>
<td>pop.</td>
<td>1 4 22</td>
<td>1 4 22</td>
<td>1 4 22</td>
</tr>
<tr>
<td>E(L)</td>
<td>0.11 0.166 0.275</td>
<td>0.377 0.395 0.431</td>
<td>1.583 1.720 1.994</td>
</tr>
<tr>
<td>σ(L)</td>
<td>0.024 0.100 0.204</td>
<td>0.042 0.065 0.105</td>
<td>0.353 0.527 0.832</td>
</tr>
</tbody>
</table>

Table 6 presents the quantiles of the distributions. For comparison the empirical quartiles, calculated from Godfrey & Neter's simulations are given as well (private communication by Godfrey, 1986). Further, Table 6 shows the quantiles P(y < L), where y relates to the population under

Table 6. Quantiles of L (in %): analytical and empirical values.

<table>
<thead>
<tr>
<th>model</th>
<th>CS1</th>
<th>CS4</th>
<th>CS10</th>
</tr>
</thead>
<tbody>
<tr>
<td>pop.</td>
<td>1 4 22</td>
<td>1 4 22</td>
<td>1 4 22</td>
</tr>
<tr>
<td>Q_1</td>
<td>0.095 0.095 0.095</td>
<td>0.334 0.334 0.334</td>
<td>1.177 1.177 1.177</td>
</tr>
<tr>
<td>Q_2</td>
<td>0.100 0.122 0.216</td>
<td>0.377 0.384 0.414</td>
<td>1.635 1.690 1.909</td>
</tr>
<tr>
<td>Q_3</td>
<td>0.119 0.203 0.403</td>
<td>0.411 0.431 0.479</td>
<td>1.891 2.036 2.392</td>
</tr>
<tr>
<td>P(y &lt; L)</td>
<td>analytical</td>
<td>empirical</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1 0.261 0.136</td>
<td>1 1 0.216</td>
<td>1 1 1</td>
</tr>
<tr>
<td></td>
<td>empirical</td>
<td>empirical</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.09 *</td>
<td>0.33 0.38</td>
<td>1.18 *</td>
</tr>
<tr>
<td></td>
<td>0.10 0.12</td>
<td>0.38 0.38</td>
<td>1.63 1.70</td>
</tr>
<tr>
<td></td>
<td>0.12 0.21</td>
<td>0.40 0.42</td>
<td>1.74 *</td>
</tr>
<tr>
<td>cov.</td>
<td>1 0.228</td>
<td>1 1 1</td>
<td>1 1</td>
</tr>
</tbody>
</table>

* number was not saved
consideration (Table 1). These probabilities are the most important characteristics of all and should exceed 95% by definition of L. The comparable empirical quantity is the coverage, taken from GODFREY & NETER (1984); it is the fraction of the calculated bounds that indeed exceeded y. (Note that the empirical figures for populations 4 and 10 were averaged.) With one exception the agreement between the analytical and the empirical values in Table 6 is quite good. The exception is Q₃ for population 1 and model CS10, where a inexplicably large difference occurs.

4. More complicated cases

Up to now, we mainly considered the case where the conditional distribution of S, given S > 0, could be described by means of a gamma density. The distribution of Z = ES¹ is easily followed, since the convolution of the gamma distributions involved was a gamma distribution again. In this section we will start the analysis of two complications, already mentioned in the foregoing parts.

(i) The gamma density is positive on the whole real axis, while S can not exceed 1 by definition. A better model of the error distribution will therefore be obtained by truncating the gamma density at 1.

(ii) Theoretical arguments and some empirical evidence indicate that a point mass at 1 may provide an even better model for the distribution of S.

As to (i), the truncated probability P(S > 1) can be reallocated in at least two obvious ways: (a) it can be smeared out over the entire interval [0,1] by dividing the original density by 1-P(S > 1), and (b) a point mass P(S > 1) at 1 can be added to the original density.

Although (b) has the advantage of meeting complication (ii), only (a) will be considered here. However, as a second extended model an additional arbitrary point mass at 1 will be added. Note that Godfrey and Neter included in their analysis some models of both kinds.

For simplicity, only the exponential distribution - as a special case of the gamma distribution - will be treated here. Starting point therefore is the (conditional) distribution Ne(a) = Γ(a,1) of S (given
S > 0) with density \( g(s) = a \exp(-as), s \geq 0 \). The two modifications indicated above will be described now in full detail; the corresponding convolutions will be derived.

I. Truncated exponential distribution

Define the constant \( c \) by

\[
(4.1) \quad c := P(0 \leq S \leq 1) = \int_0^1 g(s) \, ds = 1 - e^{-a}
\]

and the function \( g^* \) by

\[
\text{g}^*(s) = \frac{g(s)}{c}, \quad 0 \leq s \leq 1
\]

Then \( g^* \) is a density again. Denote the corresponding truncated distribution by \( N_\text{e}(a) \); see Figure 4, where the two shaded areas are equal.

![Figure 4. Truncated exponential density.](image)

For independent \( S_1, S_2 \sim N_\text{e}^*(a) \), the problem now is to derive the density \( f^* \) of \( Z := S_1 + S_2 \). Standard distribution theory gives
(4.2) \[ f^*(z) = \int_{-\infty}^{\infty} g^*(s) g^*(z-s) ds, \quad 0 \leq s \leq 1, \quad 0 \leq z-s \leq 1 \]

Since the last condition is equivalent to \( z-1 \leq s \leq z \), two cases have to be distinguished. For \( 0 \leq z \leq 1 \), (4.2) implies
\[ f^*(z) = \int_{0}^{z} g^*(s) g^*(z-s) ds = \frac{a^2}{c^2} \int_{0}^{z} e^{-az} ds = \frac{a^2}{c^2} ze^{-az} / c^2 \]
while for \( 1 \leq z \leq 2 \)
\[ f^*(z) = \int_{z-1}^{z} g^*(s) g^*(z-s) ds = \frac{a^2}{c^2} (2-z)e^{-az} / c^2 \]

By combining these results, using (4.1), we obtain
(4.3) \[ f^*(z) = \begin{cases} \frac{a^2}{c^2} ze^{-az} / (1-e^{-a})^2, & 0 \leq z \leq 1 \\ \frac{a^2}{c^2} (2-z)e^{-az} / (1-e^{-a})^2, & 1 \leq z \leq 2 \end{cases} \]

The corresponding distribution function \( F^* \) is found by integration:
(4.4) \[ F^*(z) = \begin{cases} \frac{[1-(1+az)e^{-az}]/(1-e^{-a})^2}{1-(1-2az)e^{-az}]}/(1-e^{-a})^2, & 0 \leq z \leq 1 \\ \frac{[1-(1-2az)e^{-az}]/(1-e^{-a})^2}{1-(1-2az)e^{-az}]}/(1-e^{-a})^2, & 1 \leq z \leq 2 \end{cases} \]

Of course \( F^*(z) = 0 \) for \( z < 0 \) and \( F^*(z) = 1 \) for \( z > 2 \).

II. Truncated exponential distribution with point mass at 1

Now let \( S \) have the distribution \( \text{Ne}(a) \) with probability \( \gamma \), while \( P(S=1) = 1-\gamma \); the resulting mixed distribution \( \text{Ne}^*(a,\gamma) \) is then given in detail by
(4.5) \[ \begin{cases} P(S=1) = 1-\gamma \\ g^*(s) = \gamma ae^{-as}/c, \quad 0 \leq s < 1 \end{cases} \]

Again, the distribution of \( Z := S_1 + S_2 \) will be derived for independent \( S_1, S_2 \sim \text{Ne}(a,\gamma) \).
First of all, $P(Z=2) = (1-y)^2$, since $z=2$ corresponds with $s_1 = s_2 = 1$. For $0 \leq s_1 < 1$ and $0 \leq s_2 < 1$, the analysis leading to (4.3) can be repeated, since $g^*(s) = yg^* (s)$; hence the density $f^*$ of $Z$ now is given by $f^*(z) = y^2 f^* (z)$. Finally, $0 \leq s_1 < 1$ and $s_2 = 1$ imply $z = s_1 + 1$ so that $f^*(z) = (1-y)g^*(z-1)$, while $s_1 = 1$ and $0 \leq s_2 < 1$ gives an identical expression.

Combination of the above results shows that $z$ has a mixed distribution with point mass $(1-y)^2$ at 2 and a density $f^*$ given by

\[
(4.6) \quad f^*(z) = \begin{cases} 
  \frac{y^2 a^2 z e^{-az}}{(1-e^{-a})^2}, & 0 \leq z < 1 \\
  y a e^{-az} \left[ \frac{y a (2-z)}{1-e^{-a}} + 2(1-y) e^a \right], & 1 \leq z < 2 
\end{cases}
\]

The corresponding distribution function $F^*$ equals

\[
(4.7) \quad F^*(z) = \begin{cases} 
  \frac{y^2 [1-(1+a) e^{-az}]/c^2}, & 0 \leq z < 1 \\
  y e^{-az} [y (1-2a+az)/c^2 - 2(1-y) e^a/c] + \frac{y^2 (1-2 e^{-a})/c^2 + 2y(1-y)/c,}{1 \leq z < 2} 
\end{cases}
\]

Figure 5 shows the functions $F^*$ and $F^*$ corresponding with convolutions of the distributions $N_e^*(2)$ and $N_e^*(2, 0.9)$, respectively. The dotted curve indicates the distribution function of $\Gamma(2,2)$, which is the convolution of $N_e(2)$.

5. Discussion

In this paper the analytical properties of Cox-Snell upper bounds in accounting are investigated. For a given prior model and a given population, this upper bound is a random variable $L$, the outcome of which only depends on the sampling results. The paper describes a general, theoretical method to derive the probability distribution of $L$. In this way the exact distribution of $L$ can be obtained, rendering simulation studies superfluous—at least in principle.

This method is applied in a few, relatively simple cases, where the error distribution in the population can be described by means of a
familiar density, combined with a point mass at zero. The resulting probability distributions of $L$ appeared to be in close agreement with the extensive simulation results of GODFREY & NETER (1984) with one inexplicable exception.

For more complicated error distributions, first steps are made towards a similar analysis. Unfortunately, the necessary derivations tend to be rather lengthy.

Thanks to its flexibility and nice analytical properties the Cox-Snell model may turn out to be suited for use in accounting practices. At the moment, attempts are made to implement the model for practical use.

Figure 5 Distribution functions $F^*$ and $F^+$ of the convolutions of $N^*(2)$ and $N^*(2, 0.9)$, respectively.
Acknowledgement
We are indebted to prof. J. Godfrey for communicating the additional empirical data in Table 6. We gratefully acknowledge the support of the Computer Audit Service Center of the 'Nederlandse Accountants Maatschap'.

Literature


MOORS, J.J.A. (1986), Het Bayesiaanse Cox-Snell-model bij accountantscontroles, Reeks "Ter Discussie" 86.03, Tilburg University.

IN 1986 REEDS VERSCHENEN

01 F. van der Ploeg
Monopoly Unions, Investment and Employment: Benefits of Contingent Wage Contracts

02 J. van Mier
Gewone differentievergelijkingen met niet-constante coëfficiënten en partiële differentievergelijkingen (vervolg R.T.D. no. 84.32)

03 J.J.A. Moors
Het Bayesiaanse Cox-Snell-model by accountantscontroles

04 G.J. van den Berg
Nonstationarity in job search theory

05 G.J. van den Berg
Small-sample properties of estimators of the autocorrelation coefficient

06 P. Kooreman
Huishoudproduktie en de analyse van tijdsbesteding

07 R.J. Casimir
DSS, Information systems and Management Games

08 A.J. van Reeken
De ontwikkeling van de informatiesysteemontwikkeling

09 E. Berns
Filosofie, economie en macht

10 Anna Harahczyk
The Comparative Analysis of the Social Development of Cracow, Bratislava, and Leipzig, in the period 1960-1985

11 A.J. van Reeken
Over de relatie tussen de begrippen: offer, resultaat, efficiëntie, effectiviteit, produktiviteit, rendement en kwaliteit

12 A.J. van Reeken
Groeiende Index van Informatiesysteemontwikkelmethoden

13 A.J. van Reeken
A note on Types of Information Systems

14 A.J. van Reeken
Het probleem van de Componentenanalyse in ISAC

15 A. Kapteyn, P. Kooreman, R.J.M. Willemse
Some methodological issues in the implementation of subjective poverty definitions

16 I. Woittiez
Preference Interdependence and Habit Formation in Family Labor Supply
A.J. van Reeken
A new concept for allocation of joint costs: Stepwise reduction of costs proportional to joint savings

A.J. van Reeken
Naar een andere aanpak in de systemering

J.G. de Boer, N.J.W. Greveling
Informatieplanning met behulp van referentie-informatiemodellen 1. Totstandkoming bedrijfsinformatiemodellen

J.G. de Boer, N.J.W. Greveling
Informatieplanning met behulp van referentie-informatiemodellen 2. Een methode voor informatieplanning

W. Reijnders
Direct Marketing: "Van tactiek naar strategie"

H. Gremmen
A four economy computer simulation game