Article from:

Risk Management

August 2008 – Issue No. 13

Summary of "Variance of the CTE Estimator"

John Manistre and Geoffrey Hancock

The purpose of this article is to provide a high level summary of the paper "Variance of the CTE Estimator" by B. John Manistre and Geoffrey H. Hancock that appeared in the *North American Actuarial Journal* in 2005. We also expand on some of the results.

An actuary who is responsible for estimating reserves and capital using stochastic methods must deal with a wide range of issues. Not only must the actuary worry about the underlying stochastic model, it's parameterization, data, assumptions and calculation formulae but he/she now has (or should have) the new issue of trying to quantify the precision of the estimated risk measure.

The Value-at-Risk (*VaR*) estimator (i.e., percentile or quantile value) is still often used as a risk measure. However, the Conditional Tail Expectation (*CTE*, also called Expected Shortfall or Tail-VaR) is becoming increasingly prevalent due to its desirable properties and ease of interpretation. A tool that can quantify the statistical precision of an estimated *CTE* is therefore important. That is, the sampling error in the estimate can be placed in perspective with other modeling issues, including parameter, model and assumption risk.

If all we were interested in was a regular mean then we would know what to do. We draw a sample $(x_1, x_2, ..., x_n)$ of size *n* from our model and calculate the sample average $\hat{\mu} = \frac{1}{n}\sum_i x_i$.

Statistical theory tells us three things

1. The sample average is an unbiased estimator i.e., $E[\hat{\mu}] = \mu$ the true mean. We expect to get the right answer.
2. The variance of the sample average is $VAR[\hat{\mu}] = \sigma^2 / n$ where σ^2 is the true variance.
3. If the sample size is large enough, and a few other technical conditions are satisfied, the estimator $\hat{\mu}$ has an approximate normal distribution.

The distribution's variance can be estimated from $\hat{\sigma}^2 = \frac{1}{n-1}\sum_i (x_i - \hat{\mu})^2$ and the actuary would then report the result of the work as $\hat{\mu} \pm \frac{\hat{\sigma}}{\sqrt{n}}$ giving any potential user a sense of how large the sampling error (i.e., statistical uncertainty) might be. Users can then judge whether this is large or small relative to other model sensitivities, such as a change in lapse assumptions for example, and react appropriately. In particular, users can judge whether the precision of the estimate is high enough for the intended application.

How does this process change when we start estimating Conditional Tail Expectations?

First, the bad news. There is no general set of formulas that are guaranteed to work in all circumstances. The distribution of a *CTE* estimator depends on a wide range of variables such as sample size, the actual distribution you are sampling from and the method of estimation itself.

Now, the (really) good news.

1. If the sample size being used is large enough, then there are approximate formulas analogous to those that apply for an ordinary mean. It is therefore possible to quantify the statistical precision of an estimated *CTE*.
2. There is a practical process, called a "variance verification" exercise in this article,

continued on page 38

John Manistre, FSA, FCIA, is vice president, group risk, at AEGON NV in Baltimore, Md. He can be reached at *john.manistre@aegon.com*.

Geoffrey Hancock, FSA, CERA, FCIA, is a director at Oliver Wyman in Toronto, ON. He can be reached at *geoffrey.hancock@oliverwyman.com*.

CTE Estimator ...

◗ continued from page 37

that one can execute to test (and confirm) the validity of the approximate formulas for any particular application. An example is given at the end this article.

3. Some standard variance reduction tools such as importance sampling and control variate methods can be adapted to the *CTE* problem to improve, sometimes dramatically, the precision of a *CTE* estimator for a given computational cost. These techniques will be the topic of a second article on *CTE* variance.

So what are the approximate formulas? First, we need some notation. Suppose we want to estimate the Conditional Tail Expectation of a random variable *X*, with cumulative distribution $\Pr\{X \le x\} = F(x)$ at the level α. Thus, we want to calculate the conditional expectation

$$CTE(\alpha) = E[X \mid X > q_\alpha]$$

where q_α is the α-quantile, defined as the smallest value satisfying

$$\Pr\{X > q_\alpha\} = 1 - \alpha$$

The α-quantile, is often called Value-at-Risk (*VaR*) and is used extensively in the financial management of trading risk over a fixed (usually short) time horizon.

The standard approach to this problem is to start with a random sample $(x_1, x_2, \ldots, x_n)$ of size *n* from the model and then sort the sample in descending order to obtain the order statistics $(x_{(1)} \ge x_{(2)} \ge \ldots \ge x_{(n)})$. We can then calculate the *plug-in estimators* for the required parameters by looking at the observed empirical distribution. Let $\alpha = 1 - \frac{k}{n}$ so that practical expressions for the plug-in estimators are

$$C\hat{T}E_n(\alpha) = \frac{1}{k}\sum_{i=1}^{k} x_{(i)},$$
$$V\hat{a}R_n(\alpha) = x_{(k)}.$$

In terms of this notation statistical theory has the following three things to say

1. If the sample size is large enough, and a few other very technical conditions hold, then the pair $(C\hat{T}E_n, V\hat{a}R_n)$ has an approximate multivariate normal distribution.

2. The estimator pair is asymptotically unbiased. For any finite sample size the *CTE* plug-in estimator is negatively biased i.e. $E[C\hat{T}E] < CTE$, but the bias goes to 0 as $n \to \infty$. Practical experience suggests the bias is usually much smaller than the sampling error.

3. The following approximate variance/covariance formulas are also asymptotically valid

$$VAR(C\hat{T}E_n) \approx \frac{VAR(X \mid X \ge VaR) + \alpha(CTE - VaR)^2}{n(1-\alpha)},$$
$$VAR(V\hat{a}R_n) \approx \frac{\alpha(1-\alpha)}{n[f_X(VaR)]^2},$$
$$CoV(C\hat{T}E_n, V\hat{a}R_n) \approx \frac{\alpha(CTE - VaR)}{nf_X(VaR)}.$$

The notation $f_X(VaR)$ refers to the probability density $f_X(x)$ of the random variable *X* at the point $x = VaR$.

Several comments are in order

- *CTE* is clearly easier to work with than *VaR*. If we were using *VaR* as a risk measure then we would have to find a way to estimate the probability density $f_X(VaR)$ in order to apply the asymptotic formula. This can be very difficult in practice, especially in the tails of the distribution.

- The *VaR* and *CTE* estimators are positively correlated. This makes intuitive sense.

- The variance of the *CTE* estimator has two terms. The first term is the "obvious" extension of what was happening in the first ($\alpha = 0$)

case. The origin of the second term can be seen by conditioning on the observation of the estimated $V\hat{a}R$. We can then write

$$VAR[C\hat{T}E_n] = E\{VAR[C\hat{T}E_n \mid V\hat{a}R_n]\} + VAR\{E[C\hat{T}E_n \mid V\hat{a}R_n]\}.$$

Intuitively, we can say that when we estimate the *CTE* we are estimating both the *CTE* and the $V\hat{a}R$ and uncertainty in the $V\hat{a}R$ estimate increases the uncertainty of the *CTE* estimate. This is the origin of the second term.

Simple Example— The European Put

One way to test the formulas presented above is to pick an example that is simple enough that we can get closed form expressions for all the relevant risk measures. We can then perform simulations on the model to see how well the variance estimators perform. In the formal paper we chose the example of an "in-the-money" European Put option[1] at the $\alpha = 0.95$ confidence level. Here is an edited excerpt from the paper.

To be more specific, assume the option matures in $T = 10$ years with a strike price of $X = 110$. The current stock price is $S = 100$ and assumed to follow a log normal return process with $\mu = 8\%$ and $\sigma = 15\%$. That is, the stock price at maturity is given by:

$$S(T) = S \cdot e^{\left[\mu T + \sigma \sqrt{T} \cdot Z\right]}$$

where *Z* is a standard Normal variate with mean zero and unit variance.

Using a continuous discount rate of $\delta = 6\%$, the random variable whose *CTE* we wish to calculate is then the *present value payoff function:*

$$C = e^{-\delta \cdot T} \cdot \max\left[0, X - S \cdot e^{\left(\mu T + \sigma \sqrt{T} \cdot Z\right)}\right]$$

Using spreadsheet software, we can generate $n = 1000$ samples of this variable. From this sample, we can calculate the plug-in estimators for the *CTE* and *VaR* using the formulas developed earlier. To estimate the probability density $f(VaR)$ use the estimator:

$$\hat{f}(VaR) = \frac{\xi}{\hat{F}_n^{-1}(\alpha) - \hat{F}_n^{-1}(\alpha - \xi)}$$

with $\xi = 1/100$.

We can then calculate the Formula Standard Error *(FSE)* of each estimator as

$$FSE(CTE) = \sqrt{\frac{VAR(X_{(1)},\ldots,X_{(k)}) + \alpha \cdot (C\hat{T}E - X_{(k)})^2}{n \cdot (1-\alpha)}},$$

$$FSE(VaR) = \frac{1}{\hat{f}(VaR)} \cdot \sqrt{\frac{\alpha \cdot (1-\alpha)}{n}},$$

$$C\hat{o}v(CTE, VaR) = \frac{\alpha \cdot (C\hat{T}E - X_{(k)})}{n \cdot \hat{f}(V\hat{a}R)}.$$

Table 1 shows the results of two trials (first and last) and also the results of repeating the entire simulation 1000 times. The table also shows the exact values of the *CTE* and *VaR* for this problem, which can be calculated from closed form expressions that are given in the paper.

Table 1:

Monte Carlo Simulation without Variance Reduction CTE(95%) for a 10-year European Put Option (1000 Trials), *X*=$110, *S*=$100

	$C\hat{T}E(95\%)$	$FSE(C\hat{T}E)$	$V\hat{a}R$	$FSE(V\hat{a}R)$	$C\hat{o}v(CTE,VaR)$	$\hat{f}(VaR)$
Closed Form	13.80	n/a	4.39	n/a	2.37	n/a
First Trial	13.67	1.54	5.09	1.40	1.65	0.49%
Last Trial	14.93	1.95	3.33	3.07	4.91	0.22%
Minimum	7.72	1.01	0	0.19	0.22	0.09%
Average	13.70	1.63	4.50	1.91	2.42	0.40%
Maximum	18.89	2.27	9.17	7.31	13.05	3.65%
Std Deviation	1.63	0.18	1.76	0.77	1.06	0.19%

continued on page 40

[1]A *European* put option gives the holder the right to sell the underlying asset on the maturity date for the specified strike price.

CTE Estimator ...

◗ continued from page 39

Some key takeaways from this table are:

- Any given trial provides a reasonable estimate of what would happen if the simulation were repeated, but with a sample size of $n = 1000$ there is considerable variability, especially for $V\hat{a}R$.
- The *CTE* plug-in estimator is biased below the true closed form value of 13.80 (i.e., average $C\hat{T}E$ is 13.70). However the bias is much smaller than the sampling error.
- The asymptotic variance formula for the *CTE* estimator performs quite well on average (i.e., average $FSE\left(C\hat{T}E\right)$ = empirical standard deviation of $C\hat{T}E$).
- The *VaR* plug-in estimator is biased high (average is 4.50), but again the bias is much smaller than the sampling error.
- The sample covariance for all 1000 pairs of estimators is 2.37, which is higher (lower) than the estimated covariance from the first (last) trials, but close to the mean of all covariance estimators.

A number of more realistic examples are documented in the paper using the same methodology as described above. In each case the asymptotic theory worked as expected. The examples were chosen to test a wide range of possible behaviours and practical problems facing insurers.

A More Practical Example—Variance Verification

Suppose you have a model that takes all night to run $n = 1{,}000$ scenarios. It would be impractical to repeat the run process hundreds of time in order to test the validity of the variance formulas as described above.

A more practical process for confirming the asymptotic variance formula, which we call Variance Verification, is as follows:

1. As part of model development, or in an off peak time, generate (once) a larger sample of say N = 5,000 scenarios. Let CTE_N be the estimated *CTE* based on this large sample and FSE_N the formula standard error for a given confidence level α.
2. From the large sample of size *N*, draw *m*=100 random sub-samples of size *n*=1,000 without replacement.
3. For each of the *m* sub-samples calculate a *CTE* estimate and an *FSE* estimate. Also check to see whether our best answer CTE_N lies in the approximate 95% confidence interval *CTE*±2x*FSE*. If it does we set the CI (Confidence Interval) count to 1 and 0 otherwise.
4. Use the standard deviation of the *CTE* estimates from Step 3 to check the validity of the asymptotic formula. As we will see shortly a simple adjustment needs to be made to this number before comparing it to the formula estimates.

The table below shows the results of applying the above process to an inforce portfolio of U.S. variable annuities with GMDB, GMAB and GMWB features. The book is slightly out of the money. 5000 real world (P measure) scenarios

Table 2: Variance Verification

α= 90% N = 5,000 Samples

CTE_N = 2214 FSE_N = 159

m = 100 Random Sub Samples of Size n = 1,000

	CTE	*FSE*	*CI Count*
Mean	2,111	346	94%
First	2,004	355	1
Last	2,242	353	1
Min	1,558	262	0
Max	3,120	376	0
Std Dev'n	316	30	2%

Adjusted Std Dev'n 354 *=Std Dev'n/ (1-n/N)^1/2* were generated and for each scenario the present value of guarantee benefits (claims) less the present value of related fees was captured.

The claims have been normalized so the mean $CTE(0)$ is 1,000. The fees were scaled so that $CTE(75)$ of the net (PV claims – PV fees) is zero. The example itself is for $CTE(90)$.

The first thing we note is that if the *FSE* based on the sample of size 5,000 is correct then we expect an *FSE* of about $159\sqrt{5} \approx 355$ when dealing with a sample size of 1,000. The *FSE* estimates are clearly consistent with this, but the standard deviation of 316 (of the 100 *CTE* estimates) is not. One possible explanation for this discrepancy is sampling error, but more testing shows this is not the case.

The standard deviation of the m=100 sub-samples is a biased estimate of the sampling error and it is not hard to understand why. The various sub-samples (each of size n=1,000) were all drawn from the same universe of N=5,000 so they are not independent. Because the various *CTE* estimates are using some of the same data they are positively correlated and so the set of estimates is more tightly clustered than if they were truly independent. This intuitive result is very easy to understand as $n \to N$.

It is possible to use the methods in our paper[2] to show that, in the large sample limit, the correlation of two sub-sample estimates is just $\rho = n / N$. A better estimate for the sampling error when using a sample size of 1,000 is therefore not 316, but[3]

$$\frac{316}{\sqrt{1-\frac{n}{N}}} = \frac{316}{\sqrt{1-\frac{1000}{5000}}} \approx 354 \; .$$

Our variance verification test is therefore to compare the empirical error estimate 354 with the mean formula estimate 346. We conclude that the asymptotic theory appears to be working for this model. That is, the asymptotic variance of the CTE Estimator agrees with "experiment," after adjusting for the non-independence of the sub-samples.

The CI Count result is also consistent with the idea that the formula standard errors are working. The actual count of 94 is very close to expected value of 95 (i.e., a 95% confidence interval).

Finally, it might appear from Table 2 that there is evidence of material small sample bias since the mean of the 100 sub-sample estimates is 2,111, with an apparent precision of $316/\sqrt{100} \approx 32$, which is much less than the value 2,214 obtained from the sample of size 5,000. However, this analysis is misleading, again due to the non-independence of the sub-samples.

When samples are positively correlated the variance of the sample mean is larger $VAR(\bar{x}) = \sigma^2[\rho + (1-\rho)/m]$ than it would otherwise be. A better estimate for the precision of the 2,111 number is then

$$346\sqrt{\frac{1000}{5000} + (1-\frac{1000}{5000})/100} \approx 158$$

which is not materially different from the sampling error in the 2,214 value. Thus, while there is some evidence of small sample bias in that 2,111 < 2,214, there is not enough data here to quantify it. The bias is lost in the sampling error. This is usually the case.

continued on page 42

[2] This result is not in the original paper.

[3] If we have m identically distributed, but correlated, samples from a distribution with finite mean and variance σ^2 then $E[\sum_i \frac{(x_i - \bar{x})^2}{m-1}] = \sigma^2(1-\rho)$ where ρ is the correlation between samples.

CTE Estimator ...

◗ continued from page 41

After going through a variance verification exercise a few times a practicing actuary will know whether the asymptotic formulas described here are working for his/her particular situation.

Conclusions

In the authors' experience, so far, we have yet to see a practical situation where the theory outlined here fails in a material way. Hence, we believe that practitioners can use the asymptotic formulas presented here to understand the sampling error in a given *CTE* estimate. A practitioner can go through a variance verification exercise if they want to prove to themselves, or others, that the asymptotic formulas are working in their particular situation.

Finally, we note that if a practitioner were using this tool then, on the basis of the first run of 1,000 scenarios, they would report a *CTE* estimate of 2,004±355. Is a relative sampling error of roughly $355/2{,}004 \approx 18\%$ acceptable? The answer to that question depends on the circumstances.

If 18% is not an acceptable error then what are the alternatives? One option is to increase the number of scenarios. If we increase the number of scenarios by a factor of K then the sampling error scales by a factor of $1/\sqrt{K}$. Cutting the sampling error by a factor of 4 would require us to run 16,000 scenarios.

If simply increasing the run size is impractical then there are other tools that can be used to improve the precision of the *CTE* estimator without significantly increasing the computational cost. That will be the subject of our next article. ✦