AMERICAN ACADEMY *of* ACTUARIES

Modeling Report
On the Stochastic Exclusion Test

Presented by the American Academy of Actuaries'
Modeling Subgroup of the Life Reserves Work Group

Presented to the National Association of Insurance Commissioners'
Life and Health Actuarial Task Force

Orlando, FL -- March 2008

The American Academy of Actuaries is a national organization formed in 1965 to bring together, in a single entity, actuaries of all specializations within the United States. A major purpose of the Academy is to act as a public information organization for the profession. Academy committees, task forces and work groups regularly prepare testimony and provide information to Congress and senior federal policy-makers, comment on proposed federal and state regulations, and work closely with the National Association of Insurance Commissioners and state officials on issues related to insurance, pensions and other forms of risk financing. The Academy establishes qualification standards for the actuarial profession in the United States and supports two independent boards. The Actuarial Standards Board promulgates standards of practice for the profession, and the Actuarial Board for Counseling and Discipline helps to ensure high standards of professional conduct are met. The Academy also supports the Joint Committee for the Code of Professional Conduct, which develops standards of conduct for the U.S. actuarial profession.

Modeling Subgroup

Steve Strommen F.S.A., M.A.A.A., Chair

Gordon Creber F.S.A., F.C.I.A., M.A.A.A.
Mac Smith F.S.A., F.C.I.A., M.A.A.A.
Roger Brown F.S.A., M.A.A.A.
Bruce Bohlman F.S.A., M.A.A.A.
John Froehle F.S.A., M.A.A.A.
Frans te Groen F.S.A., M.A.A.A.
Leda Debarba F.S.A., M.A.A.A.

With the valuable input of Lain Lain Kwa, A.S.A.

This report presents the results of additional modeling performed to help the Life and Health Actuarial Task Force understand the proposed stochastic exclusion test and evaluate whether it serves the intended purpose adequately.

Background and purpose of the test

In a principles-based reserving approach, stochastically generated economic scenarios are used to determine the reserve margin for risks that depend on the economic scenario. When using the CTE measure for the margin, the minimum reserve is set equal to the average of the scenario reserves under the scenarios that produce the highest scenario reserves.

For some types of contracts, the scenario reserves are not much different from one scenario to the next. Stochastic calculations serve little purpose in connection with such contracts, and their reserves could just as adequately be calculated using a reasonably conservative "deterministic scenario."

To help identify such contracts, a test for "material tail risk" was proposed in September 2007. The test involved calculating a scenario reserve under 12 deterministic scenarios. The variability of the results under those scenarios was to be used to calculate a ratio that could be used to determine whether the added effort required for full stochastic testing is justified. In light of the intended purpose of the test, this test is being re-characterized as the "stochastic exclusion test."

Major conclusions from additional modeling

The modeling discussed in the September 2007 report showed how the test ratio would vary by product type. However, LHATF requested a refinement[1] of the formula for the ratio, along with additional modeling to demonstrate that the scenarios in the test [adequately or sufficiently] covered the range of stochastic scenarios. In addition, the LRWG requested an illustration of how the test ratio might vary between a new block of business and a mature block. The additional modeling done in response to these requests is documented in the remainder of this report. The main conclusions from this work are the following:

- The limited number of scenarios adequately covers the range of stochastic scenarios for all products tested – assuming we add additional scenarios to the originally proposed 12 as provided herein. The test now proposed consists of 16 scenarios. The additional scenarios include a wider range of interest rates at later durations. These scenarios are needed to capture the tail risk of the minimum interest rate guarantee on the accumulation-style UL product that we modeled.
- The test ratio tends to be higher for a new block of business than for a mature block for the products modeled. This is probably due to the longer remaining contract life during which adverse experience can occur.
- When calculating the test ratio for a new block of business, many of the scenario reserves may be floored at zero due to the recovery of initial expenses that is built into renewal premiums. For purposes of calculating the ratio, it is important to use the scenario reserves without applying the zero floor, lest the variability by scenario be masked or eliminated.
- When calculating the test ratio for a new block of business, it is important that the scenario reserves be calculated using Gross Premium Valuation (GPV) methodology rather than Greatest Present Value of Accumulated Deficiency (GPVAD) approach. For a new block of business, the GPVAD often occurs right after the valuation date, thereby eliminating sensitivity to events that occur at later durations in a scenario. The greatest deficit occurs right after the valuation date for a new block because the recovery of initial expenses that is built into renewal premiums quickly builds up surpluses that accumulate in later durations.

The ability of the test to differentiate contracts by their level of risk was demonstrated in our earlier report and was verified in this additional modeling. With the foregoing in mind, we are comfortable proposing the test be used as a regulatory standard.

[1] The requested refinement was to change the numerator of the ratio to include only the excess of the highest scenario reserve over a base scenario reserve rather than the full range (highest minus lowest) of scenario reserves. This focuses the test on the possibility of high scenario reserves and ignores scenarios that produce reserves lower than the base scenario.

Modeling done to evaluate the test

An evaluation of the test should be based on whether the variability of the scenario reserves under the test scenarios is representative or indicative of the variability in scenario reserves under full stochastic testing. In order to determine this, we compared the results from 200 stochastic scenarios with the results of the 12 scenarios in the originally proposed test in order to see where the 12 scenario reserves in the test fell in the stochastic distribution of 200. This testing was performed for each of four different products – the same products we have used for previous reports on principles-based reserves:

- Universal Life with a secondary guarantee
- 20-year level premium term insurance
- Accumulation-style Universal Life
- Participating Whole Life

Note that the 12 scenarios are defined in a way that is intended to roughly cover the range between the 10% and 90% levels in the stochastic distribution. Since the 90% level is higher than CTE 65 for many products, the 90% level is outside the range that reserves are normally intended to cover. Given this fact, we considered it reasonable to consider the test successful if fewer than 10% of the stochastic scenarios produce scenario reserves greater than any of the test scenarios[2].

All scenarios used in this modeling begin with the December 2006 yield curve.

The following limitations of the investment strategy used in this modeling should be noted.

- The entire investment portfolio was treated as a set of fixed income investments. No equity investments were modeled. As a result, there is no variance in scenario reserves among test scenarios that differ only by the assumed equity investment returns.
- A simple investment strategy of investing free cash flow in 10-year corporate bonds with a net spread of 0.70% over 10-year Treasuries was used for all four products. Shortages of cash were handled by borrowing rather than asset sales, and the interest rate on borrowed funds was 0.80% greater than the 90-day Treasury.

It should be noted that modeling for this report was performed using prudent estimate assumptions, which are assumptions that include a margin. The modeling for our September 2007 report did not include margins in the assumptions, but it was suggested by some actuaries that margins should be included in the assumptions used for purposes of the stochastic exclusion test. There are reasonable arguments on both sides of this question.

Some arguments for including margins include:

1. The purpose of the test is to determine how the reserve depends on the economic scenario. Since margins are used when calculating reserves, they should be used in this test.
2. The inclusion of margins may make the test harder to pass for products that have risk only in the extreme tails, since margins in assumptions tend to push scenario results away from the center of a realistic distribution and towards the tails.

Some arguments against including margins include:

1. The test is designed to measure the degree of financial risk in the product, not to set the level of reserves. The degree of financial risk can be determined by using realistic assumptions in a set of scenarios to determine the degree to which results depend on the scenario.
2. If margins for investment risk are included, then this risk is essentially double counted. The margin for any risk can take either of two forms. One form is an add-on or adjustment to a deterministic assumption. The other form is to run stochastic scenarios on realistic assumptions and then average only the worst results. If deterministic margins

[2] This reasoning depends on having confidence that the 90% level in our 200 scenarios exceeds the CTE 65 level in the true distribution that would arise from a much larger number of scenarios. To gain this confidence, we carried out the following analysis using the distribution of 200 scenario results for each product. First, we calculated the CTE 65 reserve. Second, we calculated the variance of that CTE 65 estimate using the technique in the Manistre and Hancock paper in the April 2005 North American Actuarial Journal (Manistre, B. John and Geoffrey H. Hancock, "Variance of the CTE Estimator," *NAAJ*, April 2005). Finally, we determined the number of standard errors by which the 90th percentile in our distribution of 200 scenarios exceeded the estimated CTE 65 level. For term and ULSG, the 90th percentile exceeded the 65CTE by over 2.5 standard errors, a confidence level over 99%. For par whole life and accumulation UL the differences were 1.4 and 0.4 standard errors respectively, for confidence levels of 92% and 66%. Note that the stochastic distributions for the latter two products are very narrow but do have a tail associated with the minimum guaranteed interest crediting rate. That tail comes into play only in extreme scenarios and contributes to uncertainty in the CTE 65 estimate.

are used in stochastic scenarios for the same risk and one uses the tail of the stochastic distribution, one is essentially adding a stochastic margin to a deterministic margin, which some would see as double counting.

3. Products that pass the test may qualify to use simplified reserving methods that do not require the development of explicit margins. It should be unnecessary to do the work of developing margins if they are not likely to be needed for purposes of calculating the reserves under simplified methods.

Results of the modeling

The modeling led to recognition of the need for additional scenarios to cover a wider range of interest rates at later durations. The range of interest rates in a set of stochastic scenarios gets wider at later durations, and our original 12 scenarios did not reflect this much beyond the first 10 years. Four additional scenarios were added – two interest rate scenarios (high and low) each paired with two equity scenarios (high and low).

The results of modeling four different products are shown below, with results for each product on a separate page. Each page contains:

- A histogram showing the distribution of stochastic reserves, with marks showing where three of the test scenarios fall. The three scenarios marked are the minimum and maximum of the test scenarios and the base scenario.
- A table showing the scenario reserve under each of the 16 test scenarios, along with the number of stochastic scenarios greater than that scenario and the implied percentile point and CTE level on the stochastic distribution.
- The ratio that serves as the result of the test for risk.

The results show the following:

- In every case tested, the highest scenario reserve in the test scenarios exceeds the 90% point on the stochastic distribution.
- The test ratios vary substantially from product to product, and secondary guarantee universal life has the highest ratio by a wide margin. The test ratios are:

Product	Mature Block	New Block
Secondary Guarantee UL	6.8%	8.7%
20-year Level Term	1.7%	3.6%
Accumulation UL	0.8%	3.0%
Participating WL	0.2%	0.9%

Note that the test ratios are generally higher for a new block of business than for a mature block of the same product type for the products that we modeled. This makes sense because the longer the remaining life of the contracts, the longer the time period during which adverse events could occur.

Note also that the formula for the test ratio has been changed from that originally proposed. The numerator of the ratio was originally proposed to be the difference between the highest and lowest scenario reserves in the test scenarios. Since LHATF requested that only the risk of high reserves be measured, the numerator has been changed to be the difference between the highest scenario reserve and the reserve in the base scenario. Scenarios that produce a reserve lower than the base scenario reserve are ignored.

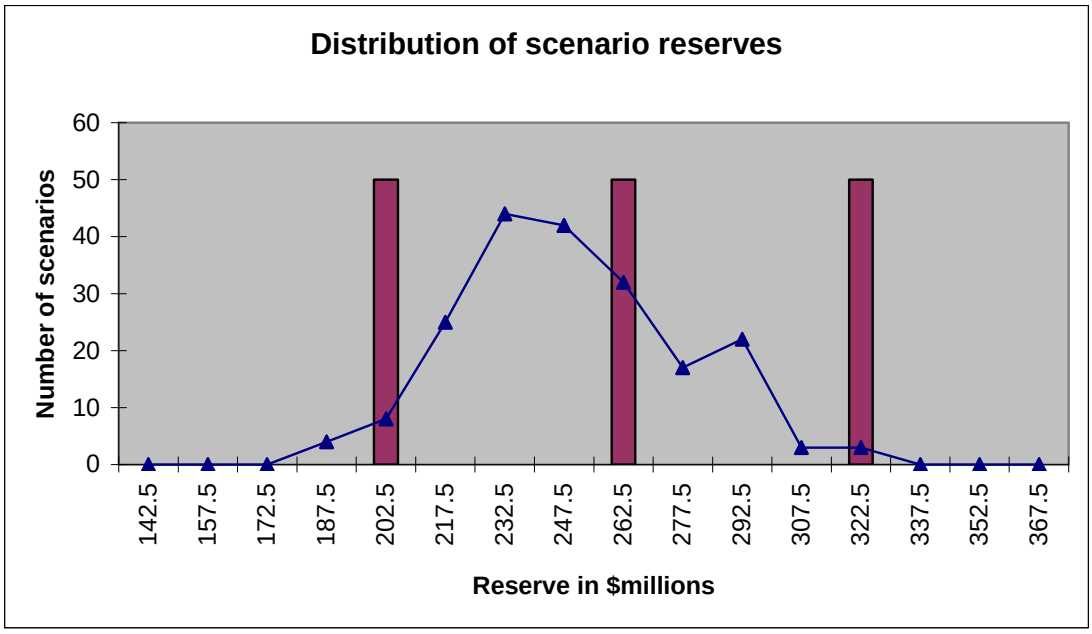
The test ratio is now defined as follows:

$$\text{Ratio} = \frac{\text{(highest scenario reserve – “base scenario” reserve)}}{\text{(base scenario reserve + present value of premiums)}}$$

The denominator of the ratio can also be expressed as the present value of benefits and expenses in the base scenario.

For this ratio to be used for the stochastic testing exclusion, LHATF must set the maximum value for the ratio that qualifies a product for the exclusion. The ratios that arose from the modeling in this report are offered as guidance in setting that maximum value, but are not intended to be a recommendation.

Results for Universal Life with a Secondary Guarantee

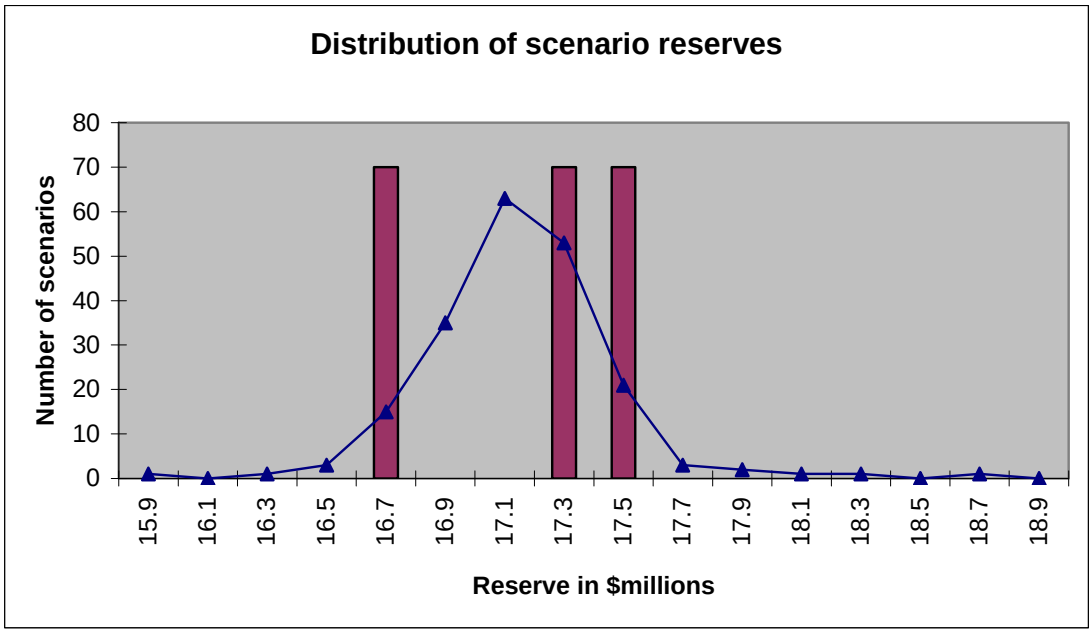


Vertical bars represent lowest and highest test scenarios, plus the base scenario

Test Scenario	Reserve	Percentile	CTE
1	198,465,897	2.5%	0.0%
2	198,465,897	2.5%	0.0%
3	308,600,745	98.5%	96.0%
4	308,600,745	98.5%	96.0%
5	225,479,043	19.5%	0.0%
6	225,479,043	19.5%	0.0%
7	271,499,368	79.5%	51.0%
8	271,499,368	79.5%	51.0%
9	259,755,772	68.5%	24.5%
10	280,855,554	83.5%	68.0%
11	259,755,772	68.5%	24.5%
12	285,421,262	86.5%	74.0%
13	229,606,885	26.0%	0.0%
14	229,606,885	26.0%	0.0%
15	287,476,732	88.0%	76.5%
16	287,476,732	88.0%	76.5%

Max reserve in test	308,600,745	A
Anticipated scenario:		
Reserve	259,755,772	B
PV future premium	457,036,643	C
Test ratio	6.8%	(A-B)/(B+C)

Results for Accumulation-style UL

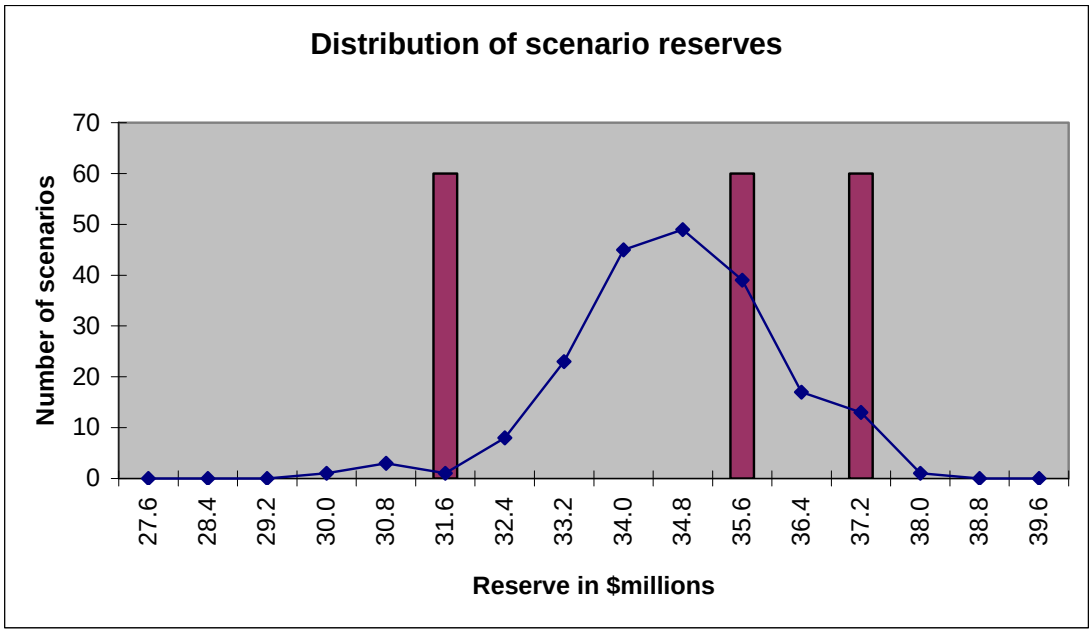


Vertical bars represent lowest and highest test scenarios, plus the base scenario

Test Scenario	Reserve	Percentile	CTE
1	16,794,876	10.0%	0.0%
2	16,794,876	10.0%	0.0%
3	16,700,045	6.0%	0.0%
4	16,700,045	6.0%	0.0%
5	17,255,257	66.5%	21.5%
6	17,255,257	66.5%	21.5%
7	17,176,163	55.0%	3.5%
8	17,176,163	55.0%	3.5%
9	17,207,921	59.0%	9.5%
10	17,003,934	28.5%	0.0%
11	17,207,921	59.0%	9.5%
12	17,007,461	29.0%	0.0%
13	16,611,342	3.0%	0.0%
14	16,611,342	3.0%	0.0%
15	17,501,263	90.5%	73.0%
16	17,501,263	90.5%	73.0%

Max reserve in test	17,501,263	A
Anticipated scenario:		
Reserve	17,207,921	B
PV future premium	19,339,016	C
Test ratio	0.8%	(A-B)/(B+C)

Results for 20-year level term

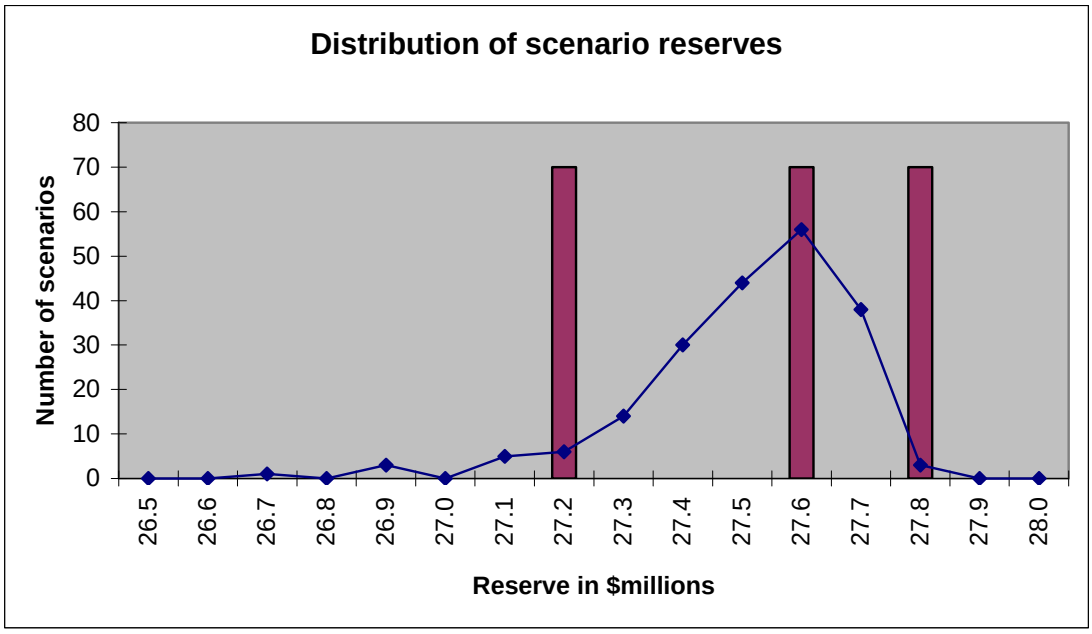


Vertical bars represent lowest and highest test scenarios, plus the base scenario

Test Scenario	Reserve	Percentile	CTE
1	31,850,259	2.5%	0.0%
2	31,850,259	2.5%	0.0%
3	37,219,686	98.0%	95.0%
4	37,219,686	98.0%	95.0%
5	32,773,544	6.0%	0.0%
6	32,773,544	6.0%	0.0%
7	36,522,839	90.0%	80.0%
8	36,522,839	90.0%	80.0%
9	35,327,726	68.5%	27.5%
10	36,935,355	94.5%	89.5%
11	35,327,726	68.5%	27.5%
12	35,840,414	80.5%	54.5%
13	35,253,546	66.5%	23.5%
14	35,253,546	66.5%	23.5%
15	35,367,388	70.5%	30.0%
16	35,367,388	70.5%	30.0%

Max reserve in test	37,219,686	A
Anticipated scenario:		
Reserve	35,327,726	B
PV future premium	76,984,879	C
Test ratio	1.7%	(A-B)/(B+C)

Results for Participating Whole Life



Vertical bars represent lowest and highest test scenarios, plus the base scenario

Test Scenario	Reserve	Percentile	CTE
1	27,192,814	6.0%	0.0%
2	27,192,814	6.0%	0.0%
3	27,718,396	94.5%	88.0%
4	27,718,396	94.5%	88.0%
5	27,579,665	58.0%	21.5%
6	27,579,665	58.0%	21.5%
7	27,566,651	56.5%	16.0%
8	27,566,651	56.5%	16.0%
9	27,589,646	62.0%	26.0%
10	27,625,083	74.0%	43.5%
11	27,589,646	62.0%	26.0%
12	27,669,356	85.0%	67.0%
13	27,071,774	2.0%	0.0%
14	27,071,774	2.0%	0.0%
15	27,739,016	97.5%	93.5%
16	27,739,016	97.5%	93.5%

Max reserve in test	27,739,016	A
Anticipated scenario:		
Reserve	27,589,646	B
PV future premium	32,550,675	C
Test ratio	0.2%	(A-B)/(B+C)

Appendix: Details of the test scenarios

The test uses a set of 16 scenarios. Given the starting yield curve on the valuation date, the scenarios are created using the Academy's stochastic scenario generator using predefined sets of random numbers.[3]

The rationale for this approach is twofold. First, the scenarios should be realistic in that they could be produced by the generator. Second, we should be able to measure in some way the likelihood of any scenario occurring.

One way to measure the likelihood of a scenario occurring is to measure the likelihood of its series of random shocks, that is, the random numbers used in the generator. Given any sequence of random numbers, their sum can be compared with a mean of zero and a standard error equal to the square root of the number of deviates in the sequence. With the mean and standard error, we can determine, in a crude way, where the sum of deviates in our sequence lies in the distribution of the sum of all such sequences.

For example, if we want a sequence that is always one standard error above average, we start with a value of 1.0 as the first deviate. The value of the n^{th} deviate is the excess of the square root of n over the square root of n-1. So the second value is $1.414 – 1 = 0.414$ and the third value is $1.732 - 1.414 = 0.318$.[4]

Generating interest rates

The Academy interest rate generator uses 3 random numbers per period. These are:

1. A random shock to the 20-year treasury rate
2. A random shock to the spread between 1-year and 20-year treasury rates
3. A random shock to the volatility

In generating the scenarios for the test, zero shocks to volatility were used.

Also, when generating scenarios for the test, upward shocks to the 20-year treasury were associated with downward shocks to the spread, making the yield curve less steep (or potentially inverted).

Generating equity returns

The Academy equity generators (C3 phase 2) use two random numbers per period. These are:

1. A random shock to make the return more or less than the mean
2. A random shock to the volatility

This potential test uses zero shocks to volatility in defined scenarios.

With that in mind, the random numbers that define the scenarios were set up as follows:

Scenario 1 – Pop up, high equity
Interest rate shocks that maintain the cumulative shock at the 90% level (1.282 standard errors).
Equity returns that maintain the cumulative equity return at the 90% level.

Scenario 2 – Pop up, low equity
Interest rate shocks that maintain the cumulative shock at the 90% level (1.282 standard errors).
Equity returns that maintain the cumulative equity return at the 10% level.

Scenario 3 – Pop down, high equity
Interest rate shocks that maintain the cumulative shock at the 10% level (1.282 standard errors).
Equity returns that maintain the cumulative equity return at the 90% level.

Scenario 4 – Pop down, low equity
Interest rate shocks that maintain the cumulative shock at the 10% level (1.282 standard errors).
Equity returns that maintain the cumulative equity return at the 10% level.

[3] Each random number is a sample from a normal distribution with mean zero and variance 1.
[4] Tables of the shocks used for the scenarios are available upon request.

Scenario 5 – Up/down, high equity, Scenario 7 – Down/up, high equity
Interest rate shocks that, for each five-year period, are consistently in the same direction. The cumulative shock for each 5-year period is at the 90% level during "up" periods and at the 10% level during "down" periods.
Equity returns that maintain the cumulative equity return at the 90% level.

Scenario 6 – Up/down, low equity, Scenario 8 – Down/up, low equity
Interest rate shocks that, for each five-year period, are consistently in the same direction. The cumulative shock for each 5-year period is at the 90% level during "up" periods and at the 10% level during "down" periods.
Equity returns that maintain the cumulative equity return at the 10% level.

Scenario 9 – Base scenario
All shocks are zero.

Scenario 10 – Inverted yield curves
Zero shocks to long term rates and equities.
Shocks to the spread between short and long rates that are consistently in the same direction for each three-year period. The shocks for the first three-year period are in the direction of reducing the spread (usually causing an inverted yield curve). Shocks for each subsequent three year period alternate in direction.

Scenario 11 – Volatile equity returns
Zero shocks to interest rates
Shocks to equity returns that are consistently in the same direction for each two-year period, and then switch directions.

Scenario 12 – Deterministic scenario for valuation
Uniform downward shocks each month for 20 years, sufficient to get down to the 80% point on the distribution of 20 year shocks. After 20 years, shocks are at a level that keeps the cumulative shock at the 80% level (or the 20% level, depending on how you look at it).

Scenario 13 – Delayed pop up, high equity
Interest rate shocks that are zero for the first 10 years, followed by 10 years of shocks each 1.414 (square root of 2) times those in the first 10 years of Scenario 1. This gives the same 20-year cumulative shock as scenario 1 but all the shock is concentrated in the second 10 years. After 20 years, the same as scenario 1.
Equity returns that maintain the cumulative equity return at the 90% level.

Scenario 14 – Delayed pop up, low equity
Interest rate shocks that are zero for the first 10 years, followed by 10 years of shocks each 1.414 (square root of 2) times those in the first 10 years of Scenario 2. This gives the same 20-year cumulative shock as scenario 2 but all the shock is concentrated in the second 10 years. After 20 years, the same as scenario 1.
Equity returns that maintain the cumulative equity return at the 10% level.

Scenario 15 – Delayed pop down, high equity
Interest rate shocks that are zero for the first 10 years, followed by 10 years of shocks each 1.414 (square root of 2) times those in the first 10 years of Scenario 3. This gives the same 20-year cumulative shock as scenario 3 but all the shock is concentrated in the second 10 years. After 20 years, the same as scenario 3.
Equity returns that maintain the cumulative equity return at the 90% level.

Scenario 16 – Delayed pop down, low equity
Interest rate shocks that are zero for the first 10 years, followed by 10 years of shocks each 1.414 (square root of 2) times those in the first 10 years of Scenario 4. This gives the same 20-year cumulative shock as scenario 4 but all the shock is concentrated in the second 10 years. After 20 years, the same as scenario 4.
Equity returns that maintain the cumulative equity return at the 10% level.

Note that the deterministic scenario for valuation (scenario 12) has not yet been recommended by the LRWG, but scenario 12 above is one possible definition that is reasonably consistent with values that have been discussed. Modeling indicates that it produces a reserve level generally between the 80^{th} and 90^{th} percentile, and a CTE level greater than 65 for three out of the four products tested.

Making scenarios available

If this test is adopted for use in VM-20, the scenarios can be made available in the same fashion as those for C-3 phase I RBC. Each year-end computer files containing the scenarios can be prepared by the Academy and made available to insurers via the Internet. A tool for generating the scenarios from any starting yield curve could also be made available.