V E R I F I C A T I O N # 3

Probability

The purpose of this verification is to confirm that the program is generating random numbers properly and applying statistics correctly. Random numbers are generated and used many different times during each simulation. If the random numbers were not being generated and applied correctly the analysis would not be statistically valid, and rational decisions could not be made based on the output. Almost all analyses performed with RocFall rely on the probabilistic nature of the program. Therefore, it is important to verify that these components are working correctly.

It is difficult to perform the verification of a random system. By definition, a random system should not be replicable; however, replication is the basis for most verification. Since the results from the program could not be replicated when the initial conditions were specified by a random distribution, all of the other verification models performed did not include a random element. This made it difficult to test the coupling of the random number generation with the algorithms that use the random numbers.

This example was designed to verify that the random number generation is being performed correctly, and that the coupling of the random number generation with the other parts of the program (e.g. the projectile algorithm) was executed correctly. Since it would be very difficult to duplicate a random procedure by hand (generating enough random samples by hand to create a statistically valid data set would be extremely time consuming), this was not done. The example was constructed so as to generate a result that could be duplicated by hand. This duplication was not achieved by reproducing each individual result, as was done with the other verification models, but rather by reproducing the result in the collective form of a random distribution.

Random Number Generation

It is important to remember that RocFall, like most computer programs, only generates pseudo-random numbers (pseudo-random implies that there is *some sort* of pattern to the numbers). The numbers are pseudo-random because the generation system uses the C runtime library's rand( ) function which produces pseudo-random output.

In order to compensate for this the number generator is "seeded" with the current time-of-day at the beginning of each simulation. This "seeding" provides a genuinely random starting point (the time) from which to begin the pseudo-random generation. This technique increases the randomness of the numbers, and should ensure that the results are sufficiently random for the purposes of rockfall analyses.

All random numbers generated in the program are sampled either from a uniform distribution or from a normal distribution. All samples from a normal distribution are created with the Box-Mueller algorithm:

$$Z_j = \mu + \sigma(\cos(2\pi S_j))\sqrt{-2(\ln(S_i))}$$

where:

Zj is a sample from the normal distribution

Si, Sj are successive samples from a uniform distribution $\in [0,1]$

μ is the mean of the normal distribution

σ is the standard deviation of the normal distribution

All samples from a uniform distributions are generated with the C runtime library's rand( ) function.

Initial Conditions

The example consists of a slope with two horizontal segments. The rocks begin by falling from a location that is slightly above the middle of the first segment. The parameters were chosen so that the location of the rock endpoints would form a normal distribution with statistical properties that could easily be determined by hand calculations.

The rocks were started at $X_0 = 0$ m, $Y_0 = 4.903325$ m. The rocks were given an initial velocity of $V_X = 5$ m/s, $V_Y = 0$ m/s.

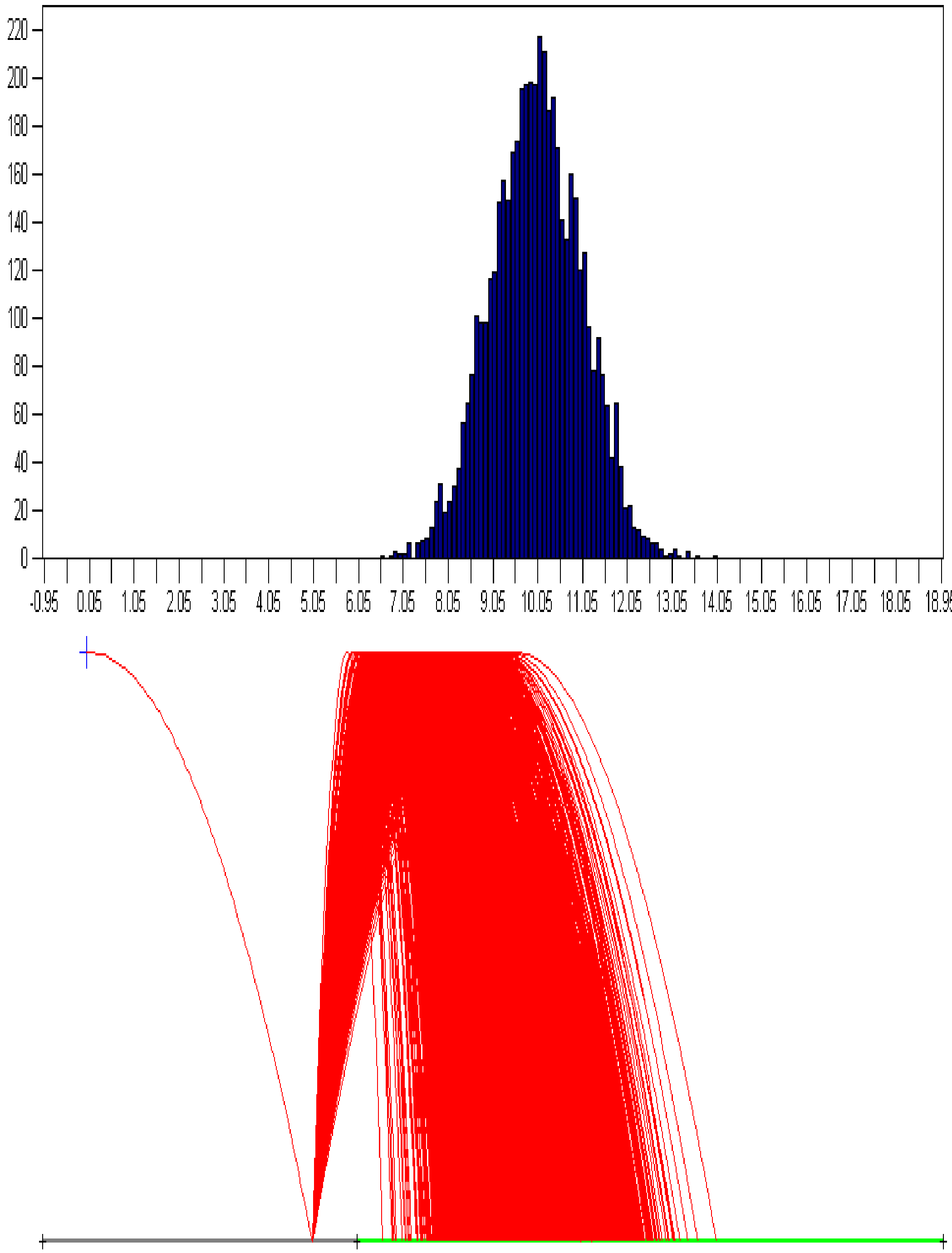
The slope geometry and coefficients of restitution are presented in the following table:

	x co-ordinate	y co-ordinate	R_N, μ	R_N, σ	R_T, μ	R_T, σ
Vertex 1	-1	0				
Segment 1			1	0	0.5	0.1
Vertex 2	6	0				
Segment 2			0	0	0	0
Vertex 3	19	0				

Table A.3.1 - Slope geometry

The coefficient of tangential restitution (R_T) along the first segment is the only parameter in the example that has any statistical variation; all other parameters are constant. R_T was given a standard deviation of 0.1 (equivalent to a variance of 0.01). The choice of $R_N = 0$ and $R_T = 0$ for the second segment was done to ensure that the rocks stopped at their respective points of impact on the second segment. R_T was chosen as the random variable because it is in the “middle” of the projectile algorithm. This was thought to be preferable to varying, say, the initial velocity of the rocks, which is only used at the beginning of the simulation. It was thought that because R_T is in the “middle” of the calculations it may be more prone to error.

The calculations were performed once, using the constant values, in order to ascertain the deterministic value for the location of the rock endpoints. The effect of the random variable (R_T) was then applied, in order to determine the expected value and standard deviation of the location of rock endpoints. The deterministic results (without the random variable) are presented first.



Number of rocks in each bin vs. distance along slope profile [m]

Figure A.3.1 - Distribution of rock endpoints

Calculations without the random variable

The rocks were started at X_0 = 0 m, Y_0 = 4.903325 m. The rocks were given an initial velocity of V_X = 5 m/s, V_Y = 0 m/s. They will fall onto the first slope segment according to (equation 4.4):

$$y = \frac{1}{2}gt^2 + V_{Y0}t + Y_0$$

Noting that $V_{YO} = 0$ and intersection with the first slope segment implies $y = 0$, equation 4.4 can be solved for t:

$$t = \sqrt{\frac{-2(Y_0)}{g}} = \sqrt{\frac{-2(4.903325)}{-9.80665}} = 1\,s$$

The intersection location and the velocity just before impact are calculated (using equations 4.3, 4.5 and 4.6):

$$x = V_{X0}t + X_0 = 5(1) + 0 = 5\ m$$
$$V_{XB} = V_{X0} = 5\ m/s$$
$$V_{YB} = V_{Y0} + gt = 0 + (-9.80665)(1) = -9.80665\ m/s$$

The pre-impact velocity is transformed into components normal and tangential to the slope segment (using equations 4.8, 4.13 and 4.14):

$$q = \frac{(Y_2 - Y_1)}{(X_2 - X_1)} = \frac{(0-0)}{(7-(-1)} = 0$$
$$\theta = \tan^{-1}(q) = \tan^{-1}(0) = 0^0$$
$$V_{NB} = (V_{YB})\cos(\theta) - (V_{XB})\sin(\theta) = (-9.80665)\cos(0) - (5)\sin(0) = -9.80665\ m/s$$
$$V_{TB} = (V_{YB})\sin(\theta) + (V_{XB})\cos(\theta) = (-9.80665)\sin(0) + (5)\cos(0) = 5\ m/s$$

The post-impact velocity is calculated by multiplying by the coefficients of restitution (using equations 4.15 and 4.16):

$$V_{NA} = R_N V_{NB} = 1.0(-9.80665) = -9.80665\ m/s$$
$$V_{TA} = R_T V_{TB} = 0.5(5) = 2.5\ m/s \qquad \text{(A.3.1)}$$

The velocities are transformed back into vertical and horizontal components (using equations 4.17 and 4.18)

$$V_{XA} = (V_{NA})\sin(\theta) + (V_{TA})\cos(\theta) = (-9.81)\sin(0) + (2.5)\cos(0) = 2.5\ m/s \ \text{(A.3.2)}$$
$$V_{YA} = (V_{TA})\sin(\theta) - (V_{NA})\cos(\theta) = (2.5)\sin(0) - (-9.80665)\cos(0) = 9.80665\ m/s$$

The rock's intersection with the second slope segment will be calculated (using equation 4.7). Noting that Y_0 = 0 and intersection with the second slope segment implies y = 0, equation 4.4 can, again, be solved for t:

$$t = \frac{-2(V_{Y0})}{g} = \frac{-2(9.80665)}{-9.80665} = 2\ s$$

The intersection location and the velocity of the rock, just prior to impact, are calculated:

$$V_{XB} = V_{X0} = 2.5\ m/s$$
$$V_{YB} = V_{Y0} + gt = 9.80665 + (-9.80665)(2) = -9.80665\ m/s$$
$$x = V_{X0}t + X_0 = 2.5(2) + 5 = 10\ m \qquad \text{(A.3.3)}$$

Incorporating the random variable

The deterministic result (x = 10 m) has been calculated above, using the constant values. The expected value and standard deviation of the rock endpoints will be calculated by applying statistical identities to the deterministic calculations.

All of the parameters remain unchanged except the value of R_T along the first slope segment. R_T is changed from a constant of 0.5 to a normally distributed random variable with a mean of 0.5 and a standard deviation of 0.1.

Before continuing, it will be useful to recall some statistical identities concerning expected value and variances. Proofs for these identities can be found in Ross (1987).

$\sigma(X) = \sqrt{\mathrm{var}(X)}$ (2.5.0 of Ross)

$E[mX + n] = mE[X] + n$ (2.5.2 of Ross)

$Var(mX + n) = m^2 Var(X)$ (2.6.2 of Ross)

where:

m and *n* are constants

σ is standard deviation

X is a random variable,

E[ ] denotes expected value

Var( ) denotes variance

The trajectory of the rock will be re-calculated, incorporating the effect of the random variable. Since R_T is the only parameter that has changed, the only equations that need to be re-calculated are equations A.3.1, A.3.2 and A.3.3.

Equation A.3.1 will be re-calculated, incorporating the random variable. Substituting $m = V_{TB}$, $n = 0$, and $X = R_T$ (expected value = 0.5, variance = 0.01) into the second and third statistical identities yields:

$$E[V_{TB}R_T + 0] = V_{TB}E[R_T] + 0$$

$$Var(V_{TB}R_T + 0) = V_{TB}^2 Var(R_T)$$

Substituting $V_{TA} = R_T V_{TB}$, $V_{TB} = 5$, $E[R_T] = 0.5$ and $Var(R_T) = 0.01$ yields:

$$E[V_{TA}] = V_{TB}E[R_T] + 0 = 5[0.5] + 0 = 2.5$$

$$Var(V_{TA}) = V_{TB}^2 Var(R_T) = 5^2(0.01) = 0.25$$

V_{TA} is now a random variable with an expected value of 2.5 and variance of 0.25.

Equation A.3.2 will now be re-calculated incorporating V_{TA} (which is now a random variable). Substituting $m = \cos(0°)$, $n = (V_{NA})\sin(\theta)$ and $X = V_{TA}$ into the second and third statistical identities yields:

$$E[\cos(0^0)V_{TA} + V_{NA}\sin(\theta)] = \cos(0^0)E[V_{TA}] + V_{NA}\sin(\theta)$$

$$Var(\cos(0^0)V_{TA} + V_{NA}\sin(\theta)) = (\cos(0^0))^2 Var[V_{TA}]$$

Substituting $V_{XA} = (V_{NA})\sin(\theta) + (V_{TA})\cos(\theta)$, $Var(V_{TA}) = 0.25$ and $(V_{NA})\sin(\theta) = 0$ ($\Theta\theta = 0$):

$$E[V_{XA}] = E[V_{TA}] + 0 = 2.5$$

$$Var(V_{XA}) = Var(V_{TA}) = 0.25$$

V_{XA} is now a random variable with an expected value of 2.5 and variance of 0.25. Since the post-impact velocity of the first trajectory (V_{XA}) is equal to the initial velocity of the next trajectory (V_{X0}), V_{X0} is also a random variable with an expected value of 2.5 and variance of 0.25.

Equation A.3.3 will now be re-calculated incorporating V_{X0} (which is now a random variable). Substituting $m = t$, $n = X_0$ and $X = V_{X0}$ yields:

$$E[tV_{X0} + X_0] = (t)E[V_{X0}] + X_0$$

$$Var(tV_{X0} + X_0] = t^2 Var(V_{X0})$$

Substituting $t = 2$, $E[V_{X0}] = 2.5$, $Var(V_{X0}) = 0.25$, $X_0 = 5$, and $x = V_{X0}t + X_0$ (equation A.3.3) yields:

$$E[x] = (2)2.5 + 5 = 10$$

$$Var(x) = 2^2(0.25) = 1$$

The standard deviation is calculated from the variance by substituting $X = x$ into the first statistical identity:

$$\sigma(x) = \sqrt{\text{var}(x)} = 1$$

Therefore, the expected value of x (the horizontal co-ordinate of the endpoints) is 10 and the variance and standard deviation of the endpoints is 1. Since x is a normally distributed random variable, the distribution of the rock endpoints should take the shape of a typical normal distribution (a bell curve) with a centre at 10. As can be seen by inspection of Figure A.3.1, the distribution is of the correct shape.

Conclusion

The same geometry and parameters were entered into RocFall and five thousand simulations were performed. The results were graphed using the "graph rock endpoints" option in RocFall. The data was extracted from the graph using the "copy raw data" option and pasted into a spreadsheet, where the statistical analysis was performed. The results from the program were compared to the manually calculated values. The results are summarised in a table:

	Hand Calculation	RocFall	Difference
Number of Samples	-	5000	-
Mean	10	9.984	0.16%
Standard Deviation	1	0.998	0.2%
Variance	1	0.996	0.4%
95% Confidence Level	-	0.0277	-
Skewness	-	0.00892	-

Table A.3.2 - Distribution of rock endpoints, comparison of results

The results from the program were very similar to the manual calculations. Given that this is a random process and *exact* answers cannot be expected, the program appears to be performing correctly; that is, the random numbers are being generated correctly and the calculations are using the random variables properly.