

CALCULO DE COTAS DE ERROR EN ACCIDENTOLOGÍA

USANDO LA TÉCNICA DE MONTE CARLO

Ernesto N. Martínez

Centro Atómico Bariloche, CNEA

Instituto Balseiro, UN Cuyo

8400 Bariloche, Argentina

emartine@cab.cnea.gov.ar

Se muestra cómo aplicar la técnica de simulación numérica llamada de Monte Carlo al tratamiento de la propagación de errores en los cálculos accidentológicos. En el caso concreto de un choque en una intersección se explican los fundamentos de la técnica de Monte Carlo, de la generación de números pseudoaleatorios, y su implementación con planillas comerciales de cálculo. Se analiza la validez y alcances de la técnica en función del número de simulaciones empleado, y se discuten los requerimientos prácticos en hardware y tiempo que plantea este método.

1. Introducción

La mayoría de las reconstrucciones de accidentes viales resultan en números (velocidades, distancias, ángulos) que están afectados de incerteza. Mientras que el cálculo de los valores numéricos en sí es un problema técnico bien entendido por el profesional que lleva a cabo la reconstrucción, la estimación de la incerteza, es decir la fijación de las *cotas de error* de los valores finales, suele constituir un desafío aún para especialistas [1, 2]. En los informes periciales es común que estas cotas o no existan, o se basen más en el ojo de buen cubero del autor del cálculo que en bases cuantitativas firmes. Así en la práctica se fijan de manera implícita cotas demasiado pequeñas, y aparecen resultados que presentan tres o cuatro cifras significativas (cuando no son las doce que es capaz de mostrar la computadora); o demasiado grandes, lo que desmerece la calidad del trabajo y le resta confiabilidad a los ojos de los usuarios finales.

Un resultado final en una reconstrucción accidentológica se deriva, a través de argumentos científicos, de sucesos más simples que son medibles. Aún cuando la cadena de razonamientos científicos sea sana, el resultado final está afectado de un error que refleja las incertezas en las mediciones elementales. El método usual para estimar cómo los errores en mediciones parciales se traducen en el error de la magnitud final, se llama de *propagación de errores*, y lo aprenden los estudiantes de ciencias exactas en los primeros años de estudio [3]. La esencia de este método es hacer un desarrollo de Taylor de la expresión matemática del resultado final en función de las magnitudes elementales, y quedarse con los términos de orden más bajo, considerando que los errores son pequeños.

Las limitaciones prácticas de este método son múltiples: no siempre se tiene una expresión matemática cerrada de la magnitud final en función de las otras; aún si se la tiene no necesariamente es simple —o posible— un desarrollo de Taylor; finalmente los errores reales no tienen porqué ser pequeños.

En los últimos años se han desarrollado, especialmente en la Física de partículas elementales, métodos que permiten estudiar *experimentalmente* los intervalos de confiabilidad de un resultado. Estos métodos se llaman *de Monte Carlo*. Debido a la cantidad de cómputo que demandan su aplicación estuvo limitada a centros de investigación; afortunadamente el uso generalizado de computadoras personales permite que hoy en día se puedan aplicar sin dificultades a los problemas accidentológicos. Este artículo explica, sobre la base de un ejemplo concreto, la forma de hacerlo: cómo aplicar la técnica de Monte Carlo a la Accidentología con los programas en uso, así como los alcances y limitaciones de este procedimiento.

2. El problema físico y su solución

La Figura 1 ilustra las posiciones finales de dos autos luego de una colisión lateral en una intersección urbana. Un Peugeot 504 con una pareja joven a bordo, embistió en el cruce a un IES Súper América manejado por una mujer joven, que no llevaba acompañantes. La "zona de impacto" que marca el croquis policial está dada por la mancha extendida de polvo fino y vidrios rotos; no se registraron marcas más

definidas del impacto. La calzada es de asfalto en buenas condiciones, seco, limpio y frío. El Peugeot ha dejado marcas de frenada de 27,9 m de largo en total, mientras que el IES no ha dejado marcas discernibles. Por lo que se sabe, ambos autos estaban en buenas condiciones mecánicas antes del accidente. No cabe duda de que el derecho de paso en la intersección le correspondía al IES. La pregunta del juez fue ¿a qué velocidad circulaba el Peugeot antes del choque?

La cadena de razonamientos que contesta esta pregunta es clara: el cálculo de la energía cinética que han disipado ambos autos por fricción luego del impacto permite hallar sus velocidades inmediatamente después de éste, mientras que la ley de conservación del momento lineal en el choque asegura que todo el momento lineal en el sentido de San Martín (horizontal en el gráfico de la Figura 1) lo aportó el Peugeot, excepto tal vez por una pequeña contribución debida a algún volantazo de último momento del IES, que no consideraremos. Finalmente, al adicionar apropiadamente la contribución del tramo de frenada preimpacto del Peugeot se obtiene su velocidad al comenzar a dejar marcas de frenada.

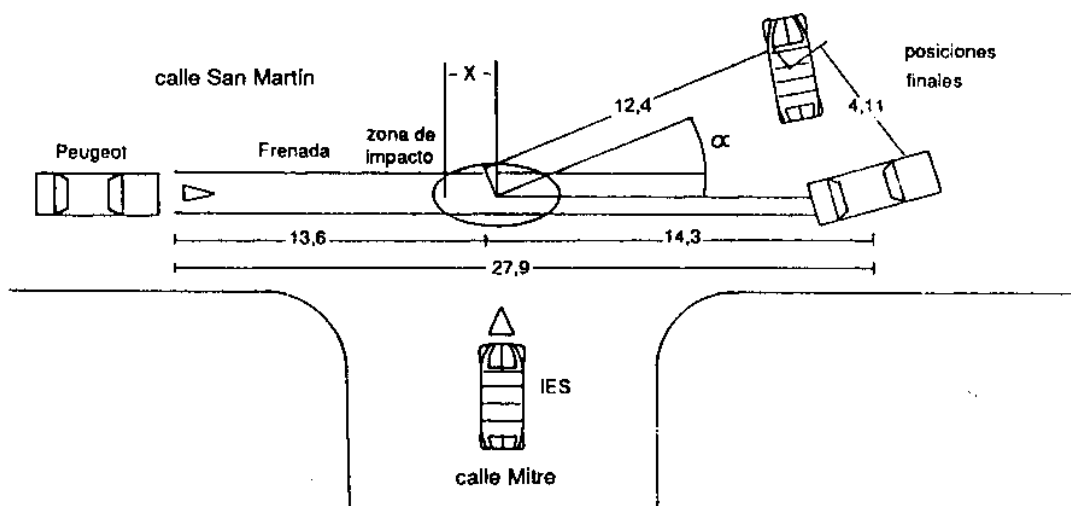


Figura 1: Croquis de las posiciones finales de ambos autos luego del choque. Las posiciones antes del choque son ilustrativas, representan sólo las orientaciones de cada auto y no sus posiciones exactas.

En principio, sólo hay que reemplazar los valores numéricos adecuados en la fórmula matemática que traduce lo dicho en palabras en el párrafo anterior para encontrar la velocidad del Peugeot, v_p . Esta velocidad inicial está dada por

$$v_p = 3.6 \sqrt{2g \left[\mu_{IES}(13.6 + x) + \left[\sqrt{\mu_P(14.3 - x)} + \frac{m_{IES}}{m_P} \frac{\sqrt{\mu_2(11.31 - x)}}{\sqrt{(11.31 - x)^2 + 4.93^2}} \right]^2 \right]} \quad (1)$$

Aquí g es la aceleración de la gravedad, m_P y m_{IES} son las masas de ambos autos, μ_P y μ_{IES} , los coeficientes de fricción para cada auto, y x , la posición real del impacto en el sentido horizontal en el gráfico. Reemplazando valores en la Expresión 1 se obtiene el valor de la velocidad del Peugeot, en kilómetros por hora, cuando comenzó a dejar sus marcas de frenada.

Sin embargo, el resultado no es único, ya que intervienen diversas fuentes de error. La más aparente es la posición del impacto, que no está bien determinada a lo largo de San Martín, donde sólo se ha marcado una zona posible. Por otro lado, los valores "adecuados" de magnitudes como las masas de los autos o los coeficientes de fricción neumáticos-asfalto son sólo promedios, y es necesario considerar las fluctuaciones estadísticamente esperables alrededor de estos promedios. Las fuentes de incerteza nombradas son las más importantes: como veremos más adelante hacen que las magnitudes respectivas varíen en el orden del 10%.

Representan los errores más importantes, pero no los únicos: cabe esperar otras fuentes de error, pero ya de menor magnitud. Por ejemplo, las longitudes medidas en el lugar por la comisión policial seguramente adolecen de errores, aunque estén bien medidas. En el caso presente, la distancia entre los dos autos en sus posiciones finales es de 4,11 m; a nadie que haya trabajado experimentalmente le sorprendería que otra pareja de policías midiera 4,08m, o tal vez 4,15m, en exactamente la misma situación. Las causas de estas

diferencias son múltiples: los puntos de referencia están mal definidos, la tensión en la cinta métrica (que ya de por sí no es un instrumento de precisión) no es siempre la misma, al leer la escala siempre se aproxima y redondea. Estos errores, llamados *estadísticos*, son absolutamente inescapables, se mida en un laboratorio de física nuclear o en la escena de un accidente; en la segunda tienden a ser mayores. También puede ocurrir que las distancias no estén bien medidas, que se hayan cometido errores gruesos: la distancia entre los autos podría aparecer como 14,11m, o tal vez como 41,10m. Sobre la magnitud de esta clase de errores poco se puede decir. En general saltan a la vista al confeccionar un gráfico a escala, ya que tienden a situar a alguno de los autos en una posición claramente imposible. En este caso, los dos autos nunca podrían estar separados por 41 metros, si se cree en las otras medidas. Estos errores a menudo se generan en la transcripción de los valores del cuaderno de notas al croquis. Volviendo a los errores estadísticos, otra fuente de error menor podría ser la apreciación de que el lugar del accidente es plano, una estimación a ojo que muy probablemente sea cierta...dentro de tal vez un uno por ciento. Así podemos buscar algunas otras fuentes de error. Sin embargo, no encontraremos muchas, y todos estos factores tendrán errores probables del orden del 1% o menores, mientras que los detallados antes son un orden de magnitud más importantes. Por lo tanto, en este artículo, y para simplificar, sólo tendré en cuenta las fuentes de error de más entidad, y despreciaré las otras; su inclusión no cambiaría apreciablemente los resultados finales. Este análisis de la importancia comparativa de las distintas fuentes de incerteza depende críticamente del caso particular; es probable que en otro caso resultara distinto. Así pues, debe tomarse como una ilustración, y no como una receta a seguir.

En suma, consideraremos aquí que los errores esperables en los resultados provienen de la indeterminación del punto de impacto en el sentido de la calle San Martín, y de las fluctuaciones en las masas de los dos vehículos y en los coeficientes de fricción asignados a la situación. Respecto a la posición transversal del punto de impacto, en el sentido de la calle Mitre, suponemos que las huellas del Peugeot y los daños en el costado izquierdo del IES lo sitúan con precisión suficiente ¿Cómo se reflejan estas fuentes de error en el resultado final? ¿Cuál es la distribución que podemos esperar estadísticamente para los valores de la velocidad inicial del Peugeot, en base a lo que conocemos sobre las fluctuaciones en las magnitudes que

entran en su determinación? La propagación de errores convencional en la Expresión 1 no es una buena opción. Por un lado, la derivación de la Expresión 1 no es una tarea que se pueda tomar a la ligera; por el otro, no tenemos ninguna garantía de que se cumplan los requisitos que permiten trincar el desarrollo de Taylor en el primer orden.

3. Método de Monte Carlo

Existe un método alternativo muy poderoso. En vez de *calcular* la propagación de errores en la Expresión 1 para la velocidad del Peugeot, se puede *generar* una cantidad de diferentes soluciones, cada una con valores distintos de los parámetros que pueden fluctuar. Estos valores distintos deben ser elegidos dentro de los límites impuestos por nuestro conocimiento del problema, es decir que deben ser todos perfectamente posibles. Cómo elegirlos dentro del rango posible es el problema: mientras que un fiscal tal vez preferiría los que conducen a valores altos de la velocidad del Peugeot, al abogado defensor se lo contrata para que busque y recomiende los que llevan a las estimaciones más bajas posibles. Está probado que los seres humanos, por imparciales que nos sintamos, no podemos elegir los valores de manera confiable. La estadística resulta mucho mejor si se deja que el azar elija los valores, que los números usados sean *aleatorios*.

Se llama “de Monte Carlo,” en alusión a las ruletas del casino de ese principado, a cualquier técnica que use números aleatorios para resolver un problema [4]. El uso más antiguo conocido de esta clase de técnicas data de 1777, cuando según se cuenta el científico francés Buffon asombró a los invitados en su salón al arrojar docenas de baguettes a ciegas por sobre el hombro. La fracción de estos panes que cortaban las juntas del piso del salón determinaba experimentalmente el número π [5]; como habían sido arrojadas al tuntún, se podía suponer que habían caído de manera debidamente aleatoria. Esta demostración tan romántica del método de la “aguja de Buffon” (usando usando agujas en vez de panes ofende socialmente menos, y funciona igual) muestra conexiones insospechadas en la ciencia, pero aparte de su popularidad como juego de salón filosófico no tuvo aplicación práctica.

Se empezó a usar el método de Monte Carlo a partir de 1944, cuando el diseño de la primera bomba atómica, y otro problema no menos arduo, la investigación operativa necesaria para montar las grandes invasiones de la Segunda Guerra Mundial, mostraron la necesidad de estos métodos. Las enormes cantidades de números aleatorios que demandan fueron provistos por las nuevas computadoras.

3.1. *Generación de números aleatorios*

Parece un contrasentido que se generen números aleatorios con una computadora, la máquina más determinista creada por el hombre. Sin embargo, en todas ellas hay disponibles rutinas que proveen de números "seudo-aleatorios": aunque son producidos de manera totalmente determinista, y por lo tanto no son aleatorios en su origen, sí lo pueden ser en la función. Generalmente los generadores de números al azar que usan los idiomas de computación son del tipo *congruencial multiplicativo*. Los números que producen forman parte de una serie determinada por un algoritmo: si se le da a la máquina una determinada "semilla", a partir de ahí la serie producida será siempre la misma. Aunque los números son deterministas, su distribución *le parece* totalmente aleatoria al usuario. En el caso del Excel de Microsoft, una planilla de cálculos muy difundida entre usuarios de PC's, un llamado a la función RAND entrega un número al azar distribuido uniformemente entre 0 y 1. Esto significa que la probabilidad de que el número caiga en un dado intervalo dentro del segmento 0,1 es proporcional a la longitud del intervalo. Por ejemplo, estos diez números fueron producidos uno tras otro por la función RAND del Excel: 0,264, 0,484, 0,585, 0,814, 0,266, 0,460, 0,465, 0,731, 0,082, 0,798. La palabra inglesa para aleatorio es *random*, que originariamente significaba "al galope", o con poco cuidado; de ella vienen los nombres de estas funciones como RAND o RANDBETWEEN(*a,b*) (un número aleatorio entre *a* y *b*).

Una advertencia para el usuario: aunque los generadores secuenciales son muy rápidos y compactos, resultan útiles si sólo se necesitan unos pocos miles de números, ya que recorren un ciclo y después empiezan a repetirse. Por esta razón, y otras más serias que no vienen al caso ahora, no alcanzan para trabajos científicos exigentes. En estos casos hay que usar otros generadores más sofisticados, para lo cual es recomendable consultar tratados especializados como *Scientific Recipes* de Press *et al.* [6] o el

capítulo 3 del segundo volumen del tratado de Donald Knuth, *Seminumerical Algorithms* [7].

3.2. Aplicación al caso presente

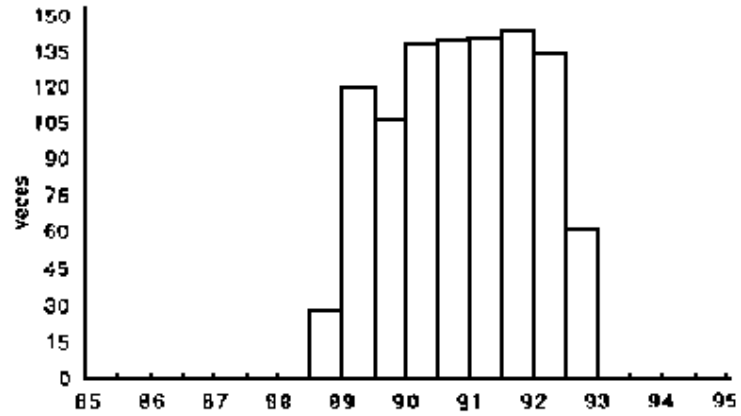
La limitada precisión que se requiere en la investigación accidentológica permite usar los generadores comerciales, sin que esto sea una recomendación para todos los que hay en existencia. Tal como ya discutimos, en la Expresión 1 figuran por lo menos cinco parámetros cuyos valores aceptan variaciones al azar. El primero es la distancia x , la posición del impacto a lo largo de San Martín. Es apropiado dejarla variar en una zona de 3 m, de forma uniformemente distribuida. Eso quiere decir que x va de $-1,5\text{m}$ a $+1,5\text{m}$, y que en términos de la función RAND esta variable aleatoria podría definirse por la instrucción:

$$x = \text{RAND()} * 3 - 1.5$$

El valor particular de esta variable determina los valores de las tres distancias del problema, así como el del ángulo α .

La Figura 2 muestra el histograma de 1000 valores de la velocidad inicial del Peugeot, obtenida de la Expresión 1, con valores de x variando al azar según la expresión de arriba, mientras los otros parámetros permanecen fijos. Los valores extremos de la velocidad son 88,92 y 92,70 km/h, con un valor medio de 90,87 km/h. La desviación estándar (raíz cuadrada del segundo momento centrado) es $\sigma = 1,07$ km/h. Estos valores son propios de esta *realización*, de este conjunto particular de 1000 simulaciones, y otra corrida daría otro conjunto de valores, con otros extremos, otro valor medio, y otra desviación estándar. Más adelante se analizarán cómo varían estos parámetros de realización a realización.

Figura 2. Histograma de 1000 valores de la velocidad inicial del Peugeot, determinados por la variación en el punto de impacto.



Las masas de ambos autos también pueden variar, de forma independiente. Considerando los pesos de los autos en orden de marcha [8], más los probables de sus ocupantes, las fluctuaciones aleatorias en las masas de los autos pueden modelarse por:

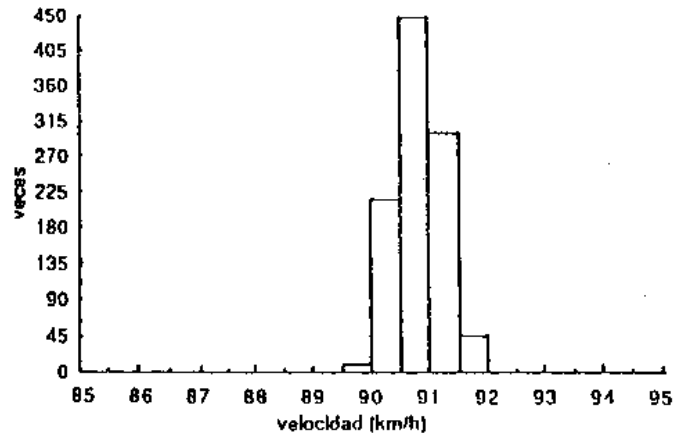
$$m_P = \text{RAND}() * 60 + 1267$$

$$m_{IES} = \text{RAND}() * 40 + 550$$

Esto es, consideramos que los autos pesan por lo menos 1267 Kg y 550 Kg, y a esos pesos se les puede sumar cualquier cosa entre cero y 60 o 40 kilos, según el auto. Las masas están expresadas en kilogramos, pero no estoy usando unidades en la planilla de cálculos. En el caso de que se calcule con un programa como MathCad es conveniente usar las magnitudes con sus unidades, una costumbre que ayuda a evitar errores de cálculo al imponer la consistencia dimensional. De hecho, en este caso especial es irrelevante qué unidades se usan para las masas, ya que en la Expresión 1 aparece sólo su cociente.

La Figura 3 muestra el histograma de 1000 valores de la velocidad variando sólo estas masas. Los valores extremos son 89,93 y 91,77 km/h, con una media de 90,83 km/h y $\sigma = 0,40$ km/h. Este último valor deja en claro que los errores en las masas son de importancia menor que los provenientes del punto de impacto.

Figura 3. Histograma de 1000 valores de la velocidad inicial del Peugeot, determinados por la variación de las masas de los autos.

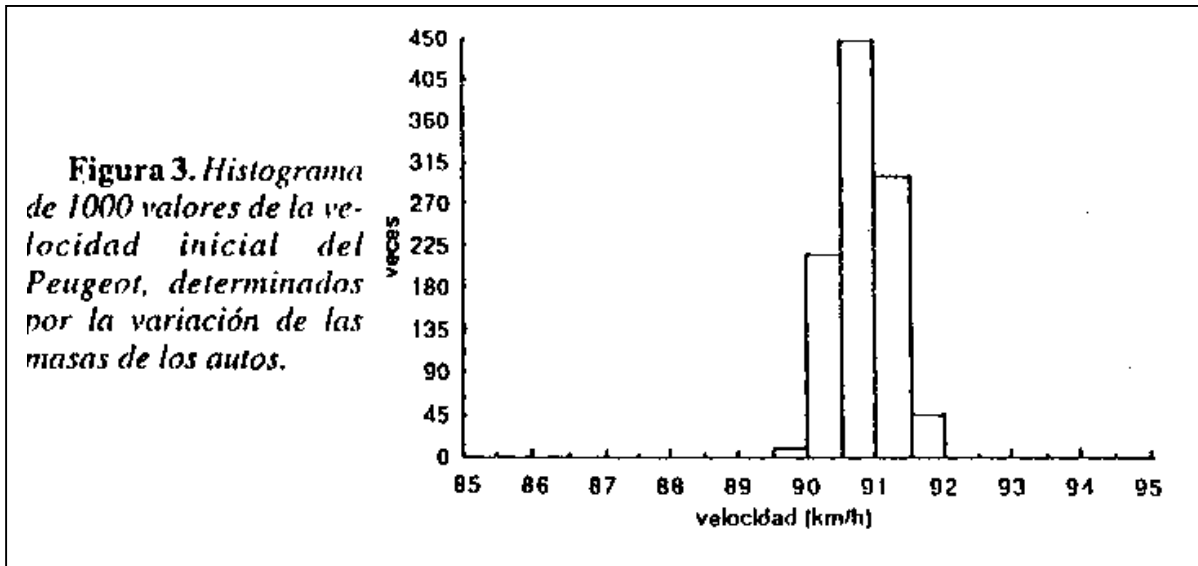


La variación en los coeficientes de fricción efectivos de cada auto reviste especial importancia. En las circunstancias del accidente un promedio razonable es 0,8 [9,10], pero una variación en hasta 0,1 para arriba y abajo también es razonable:

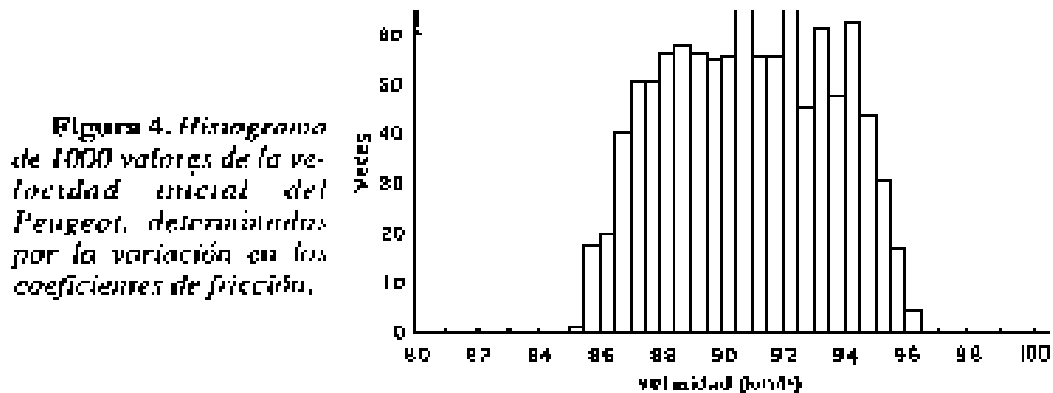
$$m_p = \text{RAND}() * 0.2 + 0.7$$

$$\mu_{\text{IES}} = \text{RAND}() * 0.2 + 0.7$$

Nótese que, aunque las expresiones son iguales para ambos coeficientes de fricción, m_p y m_{IES} , los valores numéricos que toman son distintos e independientes, ya que la función RAND proporciona un número aleatorio distinto *cada vez* que se la invoca.



En la Figura 4 se ve el histograma de los 1000 valores determinados variando μ y μ_{IHS} , con los otros parámetros quietos. Nótese el cambio de la escala horizontal, exigido por el rango mayor en la variación de la velocidad: sus valores extremos ahora son 85,44 y 96,21 km/h. El valor medio es de 90,88 km/h, y la desviación típica σ es de 2,69 km/h. Esta fuente de error es claramente la más importante.

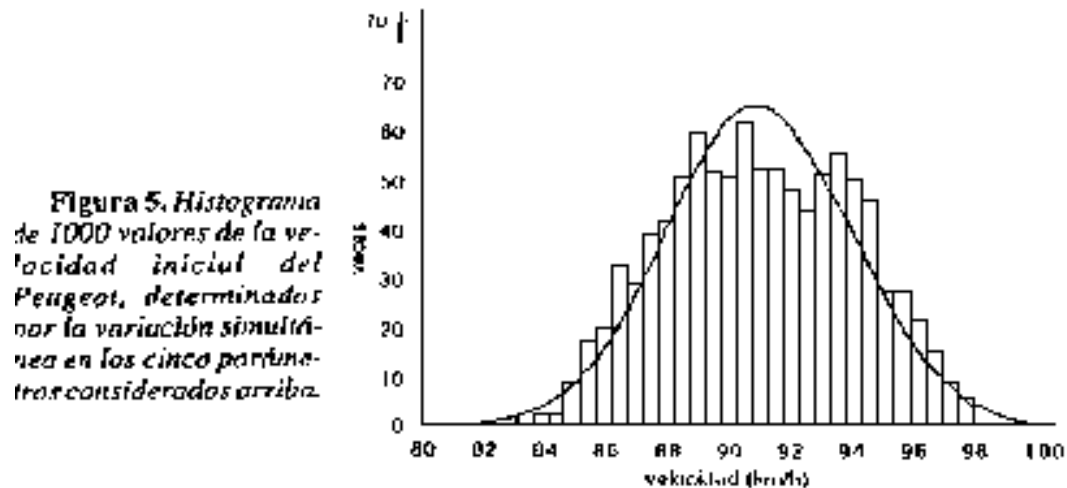


Por último, la Figura 5 muestra el histograma de 1000 valores de la velocidad en los cuales han variado los cinco parámetros simultáneamente. Representa por lo tanto un resultado del modelo de variación total. Aquí los valores extremos son 83,71 y 98,21 km/h. El valor medio es 90,78 km/h, y $\sigma = 2,98$ km/h. Como era de esperarse, la distribución de velocidades en la cual varían todos los factores es más ruidosa, tiene mayor dispersión que cuando varían sólo algunos de los parámetros. La desviación estándar total

resulta ser muy aproximadamente la raíz cuadrada de la suma de los cuadrados de las desviaciones parciales:

$$s = 2.98 \approx \sqrt{1.07^2 + 0.40^2 + 2.69^2} = 2.92$$

un resultado típico para distribuciones gaussianas.



La Figura 5, que muestra superpuesta al histograma una distribución gaussiana (o *normal*) con los parámetros adecuados ($N = 1000$, $\langle v \rangle = 90.78$ km/h, $\sigma = 2.98$ km/h), permite ver que el histograma ya ha adoptado una forma de campana típicamente gaussiana que no mostraban los anteriores. Esto es una consecuencia del efecto simultáneo de varias fuentes de fluctuación. Aunque ninguna de ellas obedece de por sí una distribución gaussiana, su combinación hace que los valores extremos, que exigen valores extremos *simultáneos* en todas las variables, sean mucho menos probables que los moderados del centro. Este acampanamiento de la distribución de valores de la velocidad muestra el Teorema Central del Límite en acción: bajo condiciones muy generales (pero no universales), la combinación de numerosas fuentes de variación da origen a una distribución gaussiana de valores finales.

Una segunda mirada más crítica permite ver que, aunque la distribución se parece

satisfactoriamente a una gaussiana, no lo es exactamente, ya que su centro es demasiado bajo y sus alas demasiado altas. En efecto, el *coeficiente de curtosis* de esta distribución, definido como el cociente entre el momento de cuarto orden y el cuadrado del de segundo orden (ambos centrados) da 2,18, mientras que el de una auténtica distribución gaussiana debería ser 3 [11]. La distribución de velocidades es *platicúrtica* a pesar de ser de rango limitado. Esta propiedad proviene de que una de las fuentes de error, la variación en los coeficientes de fricción, predomina netamente sobre las otras. Como su forma es rectangular, la distribución final conserva un núcleo rectangular que las demás variaciones no han podido diluir, y confiere el carácter platicúrtico a la distribución. En cuanto a la leve *asimetría* que muestra esta distribución, proviene de la raíz cuadrada que domina la Expresión1, pero es tan pequeña, de 0,0014, que no resulta relevante.

3.3. *Confiabilidad de los resultados*

La simulación numérica brinda una manera rápida y muy poderosa de determinar rangos de variación, e inclusive formas de distribución, de resultados que pueden no ser accesibles a cálculos convencionales de propagación de errores. Ahora bien: ¿hasta dónde se pueden creer estos resultados *experimentales*?

Para empezar, la computadora más poderosa del mundo no es mejor que quien la programe. Esto lo expresan los programadores en las siglas inglesas GIGO, que se podrían traducir como EBSA: si entra basura, sale basura. Si se parte de un modelo físico errado para la colisión, o no se estiman adecuadamente los rangos de variación de los parámetros, o se desprecia la variación en los parámetros realmente relevantes del problema, los resultados no valdrán nada. Pero estas reservas valen también para cualquier cálculo analítico tradicional. Hay que considerar los aspectos intrínsecos de la técnica de Monte Carlo.

El primer aspecto que debe cuidarse ya fue mencionado: los generadores de números aleatorios incluidos en los paquetes comerciales son pobres. Son adecuados para generar algunos miles de números, pero si se quiere algo más ambicioso se debe recurrir a rutinas mejores [6, 7]. El segundo aspecto es

inherente al carácter estadístico del método: si cada realización nos provee de N valores distintos de la velocidad del Peugeot, y estos N valores dan su propio valor medio y desviación, ¿a cuál realización habría que creerle?

La respuesta la da la estadística: las fluctuaciones en estos valores son inversamente proporcionales a la raíz cuadrada de N . Esto significa que al hacer simulaciones con N cada vez mayores, las cifras en los resultados que dependen del conjunto, sea el promedio o la desviación estándar, se van quedando fijas desde la izquierda. Es un problema de muestreo, presente en todas las encuestas: en cada simulación obtenemos sólo un número finito, N , de respuestas, sean velocidades posibles o intenciones de voto. Si son pocas, los parámetros de interés (valores extremos, promedio y desviación estándar) están muy sujetos a las variaciones del azar. Al aumentar el tamaño de la muestra los caprichos del azar tienden a compensarse, y los parámetros calculados se tornan más confiables.

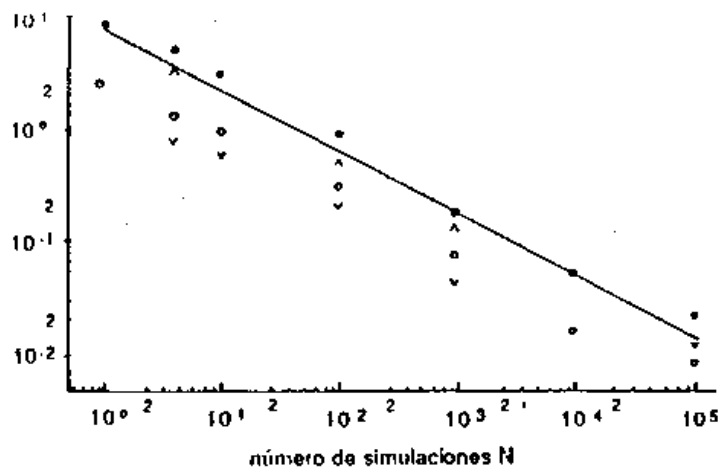
Para mostrar esta tendencia, he hecho 10 realizaciones distintas, cada una de N simulaciones, y he repetido las simulaciones para N de 1 a 100000. He calculado el promedio y la desviación estándar para cada realización. La Figura 6 muestra, en escalas logarítmicas, la diferencia entre los *valores extremos* en cada grupo de 10 realizaciones, del promedio y de la desviación estándar.

Esta diferencia entre los valores extremos es la medida más radical posible de las fluctuaciones, el rango total de variación, no el estadísticamente significativo que, por supuesto, es menor. La recta, dada por uno la sobre raíz cuadrada de N , sirve de comparación. Se ve claramente que los rangos de dispersión de ambas magnitudes son efectivamente inversamente proporcionales a la raíz de N , y que para aumentar en un factor 10 la precisión estadística de los resultados, es decir para asegurar una cifra decimal más, hay que calcular 100 veces más. Las técnicas de Monte Carlo son de fuerza bruta: se basan en el poder de cálculo rápido y barato para reemplazar al más artesanal y caro.

Pero no porque el cálculo sea barato hay que derrocharlo. La pregunta en este caso es ¿qué precisión es

necesaria en el problema que se está tratando? Técnicamente, se especifica la desviación típica de un resultado con una o dos cifras significativas, nunca con más. ¿Cuánto debe ser N para obtener esta precisión? La Figura 6 muestra que con $N = 100$ ya se ha fijado la primera cifra de σ , igual a 3, pues todos los valores caen cómodamente entre 2,8 y 3,3. Con $N = 10000$ ya sobra para tener dos cifras seguras, $\sigma = 3,0$, pues los valores caen entre 2,97 y 3,04. Las 1000 simulaciones usadas en las Figuras 2 a 5 parecen ser un buen número, y aseguran dos cifras significativas si se usa el criterio estadísticamente más realista de la dispersión cuadrática media, en vez del rango total de variación.

Figura 6. Rangos totales de variación del promedio (\bullet) y de la desviación típica (σ), y desviaciones típicas del promedio (Δ) y de la desviación típica (∇), para 10 realizaciones, de N simulaciones cada una, en función de N . La recta sirve de referencia.



Personalmente, creo que usar dos cifras significativas en la desviación estándar de un cálculo accidentológico (o de cualquier otro tipo) es hilar demasiado fino, y que basta con una. En efecto, usar dos cifras significativas en la desviación estándar equivale a asegurar que conocemos el efecto de los errores mejor que el 10%, y eso me parece exagerado. Así pues, es razonable decir que la desviación estándar de los resultados es 3 km/h. Consecuentemente, el promedio debería redondearse a este grado de precisión, con lo cual quedaría 91 km/h. Desde un punto de vista estadístico, entonces, se puede expresar el resultado diciendo que la velocidad del Peugeot al comenzar a dejar marcas de frenada era de 91 ± 3 km/h.

Estadísticamente, esto significa que el 68% de las soluciones están comprendidas dentro de este rango. Para propósitos forenses, el 32% que queda afuera es demasiado (en realidad, lo que puede preocupar a la

defensa, o al perito que trata de no perjudicar a un imputado, es sólo la mitad, el 16% que queda por *debajo* del rango; del 16% que cae por *arriba* la defensa prefiere ni hablar). Parece más razonable, en este ámbito, aumentar el rango de error, por ejemplo a 3σ . Este intervalo comprende el 99,7% de los casos. El uso de un intervalo de error de 3σ en cada sentido daría un resultado final de 91 ± 9 km/h, preferible a mi juicio. Y como ya el 1 del 91 resulta algo pedante comparado con el error de 9 km/h, sería preferible redondear y decirle al juez que la velocidad más probable era de 90 km/h, y seguramente estaba entre 80 y 100 km/h. Según mi experiencia, muy raramente se justifica una precisión mayor que un 10% en una reconstrucción accidentológica.

3.4 Máquina y tiempo

Hemos visto que en un caso como el discutido, con fuentes de error típicas, alcanza con unas 1000 simulaciones para obtener una muy buena idea, y una medida cuantitativa, del error probable en el cálculo a través de la técnica de Monte Carlo. La pregunta ahora es cuánto cuesta, en hardware y tiempo, hacer estas simulaciones.

“Ahora” en la frase anterior se refería a mediados de 1993, cuando escribí la versión original de este artículo. Resulta interesante desde un punto arqueológico comparar mis respuestas de entonces con lo que sucede diez años más tarde, al preparar esta revisión.

Sigue lo escrito en 1993. Respecto al primer punto, para hacer los cálculos que he presentado he usado una máquina poco sofisticada: una PC AT 286 de 16 Mhz, con coprocesador matemático. La falta del coprocesador haría los cálculos algo más lentos. Los cálculos que se muestran en las Figuras 2 a 5 están hechos con una planilla de cálculos comercial que da números al azar. Cada conjunto de 1000 simulaciones tardaba 24 segundos. La planilla de cálculos es lenta, pero es fácil de usar, y permite analizar estadísticamente y graficar directamente los resultados, una gran ventaja en el trabajo práctico. Para hacer los cálculos de la figura 6, la planilla de cálculos me resultaba demasiado lenta, además de que no podía manejar 100000 simulaciones, y escribí un programa de Quickbasic que producía y analizaba

estadísticamente 1000 simulaciones en 2,1 segundos, y en 1,87 segundos si estaba compilado. Las 100000 simulaciones tardaban 185 segundos. Al escribir programas de este tipo es muy importante, si se quiere acelerar la marcha, evitar operaciones dentro del ciclo de repetición: cualquier ahorro se multiplica por 100000. Por ejemplo, la multiplicación por $3.6 \sqrt{2g}$ en las velocidades se debe hacer fuera del ciclo.

En 2003. Mi máquina ahora es una Pentium, nadie se compra un coprocesador, y las realizaciones de 1000 simulaciones son, por lo que puedo ver, instantáneas. El idioma de computación Quickbasic ya no viene incluido cuando uno compra una PC, una falta compensada por diversos programas que ya traen paquetes armados para hacer simulaciones de Monte Carlo, por ejemplo el Analysis Tool Pack de Excel, el MathCad o el Matlab.

4. Conclusiones

El método de simulación de Monte Carlo presenta una manera eficaz y práctica de tratar la propagación de errores en los cálculos de la reconstrucción de accidentes. Hoy en día los medios técnicos para llevarla a cabo están al alcance de cualquier profesional en Accidentología. Cualquier computadora personal, y casi cualquier planilla comercial de cálculos, permiten la aplicación de este método aún por parte de personas no avezadas en computación. La aplicación generalizada de esta técnica llevaría a estimaciones más realistas y acertadas de los rangos de error probable, y por lo tanto a veredictos más atinados y confiables. Probablemente una simulación de Monte Carlo sea demasiado engorrosa para llevarla a cabo en cada reconstrucción, pero esto no es necesario. Basta con aplicarla a cada situación nueva y usar sus resultados para estimar los errores finales y así ir adquiriendo una intuición bien informada sobre la validez de nuestros resultados.

Referencias

[1] Willem I. Zijp. *Treatment of measurements and uncertainties*. Informe técnico ECN-194, Netherlands Energy Research Foundation ECN, Petten, 1987.

- [2] F. James. *Determining the statistical significance of experimental results*. Informe técnico DD/81/ 02, CERN Data Handling Division, Geneva, 1981.
- [3] H. J. J. Braddick. *The physics of experimental method*. Chapman & Hall, London, 1963.
- [4] F. James. *Monte Carlo theory and practice*. Informe técnico DD/ 80/6, CERN Data Handling Division, Geneva, 1980.
- [5] G. L. Leclerc de Buffon. *Essai d'arithmetique morale*. París, 1777.
- [6] W. H. Press, V.B.T. Flannery, S.A. Teukolsky, y W. T. Vetterling. *Numerical recipes: The art of scientific computing*. Cambridge University Press, Cambridge, 1986.
- [7] Donald E. Knuth. *The art of computer programming*. Addison-Wesley, Reading, Massachusetts, 1981.
- [8] Carlos A. Hirz, editor. *El cronista automotor*. Ediciones El Cronista Comercial, Buenos Aires, 1988.
- [9] Charles Y. Warner, G. C. Smith, M. B. James y Geoff J. Germane. *Friction applications in accident reconstruction*. Informe técnico 830612, Society of Automotive Engineers, Warrendale, 1983.
- [10] James C. Collins. *Accident reconstruction*. Charles C. Thomas Publisher, Springfield, 1979.
- [11] Murray R. Spiegel. *Estadística*. Mac Graw-Hill, Madrid, 1991.