Taller de Probabilidad y Simulación

Probabilidad I

Departamento de Matemáticas UAM, curso 2008-2009

Pablo Fernández Gallardo (`pablo.fernandez@uam.es`)

1. Lista de posibles temas de trabajo

En lo que sigue exponemos una serie de temas que pueden ser elegidos como trabajos de fin de curso del Laboratorio de Simulación (Probabilidad I, curso 2008-09). Cabe también la posibilidad de desarrollar alguna cuestión que no aparezca en esta lista (previa consulta con el profesor).

En cada uno de ellos se da una breve descripción que incluye: el objetivo, el modelo y un esquema de trabajo que recoge, como orientación, algunas posibles preguntas al respecto del problema planteado. Las descripciones son en ocasiones expresamente vagas, porque uno de los objetivos de estos trabajos es, justamente, plantearse las preguntas adecuadas. Y, por tanto, cualquier variación o ampliación interesante sobre lo aquí escrito será bienvenida.

1. Simulación del juego del tenis

- **Objetivo.** Simulación del juego del tenis para analizar la relación entre la probabilidad de ganar puntos sueltos y la probabilidad de ganar partidos. Análisis del efecto que sobre esta relación tienen posibles cambios de reglas (más sets de menos juegos, por ejemplo).
- **Modelo.** Dos jugadores: A y B. El jugador A gana cada punto con probabilidad p y B lo gana con probabilidad $(1-p)$. Los puntos se ganan o pierden independientemente unos de otros.
- **Esquema de trabajo.** Determinar empíricamente la relación entre p y la probabilidad de ganar el partido y el número medio de puntos de que consta cada partido. Repetir el análisis con reglas distintas: de cuatro sets de cuatro juegos, por ejemplo. O quizás como en el juego del ping-pong: partidos a un cierto número de puntos (sin sets).

  Repetir el análisis con una variación (para hacerlo más realista): ahora A gana los puntos "normales" con probabilidad p y los puntos "comprometidos" (decisivos) con otra probabilidad p^*.

2. En el casino

- **Objetivo.** Simulación de ruleta americana y europea para analizar las consecuencias de distintas estrategias de apuestas a Rojo y Negro
- **Modelo.** Las distintas casillas de la ruleta son igualmente probables. En la ruleta europea están los números del 1 al 36 (la mitad rojos, la mitad negros), además del 0. En la americana, además, hay un 00. Los números van saliendo de manera independiente.
- **Esquema de trabajo.** Analizamos la cuestión conocida como la *ruina del jugador*. Partimos de una cierta fortuna inicial n. Apostamos siembre al rojo 1 euro. Si sale rojo, ganamos uno más, y si no sale rojo, perdemos el euro apostado. Jugamos hasta que, o bien alcancemos una fortuna N, o bien nos arruinemos (la fortuna inicial está entre 0 y N). Se trata de obtener empíricamente la gráfica: probabilidad de arruinarse frente a p (la probabilidad de ganar cada apuesta). Para distintos valores de p (para $p = 1/2$, para el p de la ruleta europea, el de la americana, y otros menores). Calcular duración media de las partidas. Por ejemplo, para valores $N = 40$, $n = 20$.

  Variación: en la misma situación de antes (fortuna inicial n, objetivo N), queremos maximizar la probabilidad de alcanzar el objetivo. La probabilidad p puede ser 1/2 o menor. Se trata de comparar diversas estrategias de apuestas, como las siguientes:

- Apostar una cantidad fija (no necesariamente 1) siempre a Rojo.
- Apostar siempre toda la fortuna disponible.
- *Bold play*: en cada jugada apostamos todo el capital disponible, salvo que una apuesta menor sea suficiente para poder alcanzar el objetivo en la siguiente tirada.

Nota: convendría hacer simulaciones "realistas", en las que el número de apuestas y el montante de éstas tengan restricciones.

3. Tiempo de espera de autobuses

- **Objetivo.** Analizar la influencia de la dispersión de los intervalos entre llegadas sobre el tiempo medio de espera.
- **Modelo.** La unidad de tiempo es el minuto. El intervalo de tiempo A entre un autobús y el siguiente es una variable aleatoria finita positiva (que toma valores enteros) con media m minutos, mientras que la llegada V del viajero a la parada se modeliza con una variable aleatoria uniforme entre los minutos de un amplio intervalo de tiempo.
- **Esquema de trabajo.** Estimar el tiempo medio de espera del viajero, y obtener empíricamente una relación entre la media y la varianza de A y la media de V. Obtener empíricamente la distribución de probabilidad del número de autobuses que se observan en un intervalo arbitrario de tiempo de amplitud $3m$.

4. Cartera de inversiones

- **Objetivo.** Analizar el resultado de una estrategia de inversión en acciones en función de la correlación
- **Modelo.** Normal multidimensional.
- **Esquema de trabajo.** Nuestra cartera consta de α unidades del activo A, y β del activo B. Su precio hoy es conocido. Queremos analizar sus posibles precios en un cierto horizonte temporal T. Modelamos las variaciones porcentuales de cada activo con una variable normal con ciertos parámetros. Conjuntamente, las variaciones siguen una normal multidimensional con correlación ρ. Obtener el histograma (junto con los estadísticos adecuados: media, desviación típica, algún percentil) de los posibles valores de la cartera en función de ρ.

  Variación: incorporar a nuestra cartera *derivados* sobre el precio de las acciones. Es decir, instrumentos financieros cuyo pago depende del precio de la acción en T. Como, por ejemplo, una *call*, cuyo pago es $\max(X - K, 0)$, donde X es la cotización en tiempo T de la acción y K es el *strike* (un cierto número).

5. Patrones en lanzamientos de cara y cruz

- **Objetivo.** Entender cómo la estructura de los distintos patrones influye en sus tiempos de aparición o a la hora de "competir" entre ellos.
- **Modelo.** Lanzamientos independientes de moneda (quizás cargada).
- **Esquema de trabajo.** Simulamos del lanzamiento de una moneda hasta que sale un patrón (una cierta lista) de caras y cruces predeterminado. Calcular tiempos medios de aparición. Ahora ponemos a competir patrones (por ejemplo, de longitud 3 ó 4) con distintos tiempos medios de aparición.

6. Ley del arco seno de las rachas en el camino aleatorio

- **Objetivo.** Entender las sutilezas de un objeto tan aparentemente sencillo como el camino aleatorio (suma de sucesivos $+1$ y -1 con lanzamientos de moneda).
- **Modelo.** Lanzamientos independientes de una moneda (equilibrada).
- **Esquema de trabajo.** Lanzamos una moneda un número N determinado de veces; en cada lanzamiento, si sale cara, ganamos 1, y si sale cruz perdemos 1. Una racha es una serie consecutiva de lanzamientos para los que nuestra fortuna está por encima (o por debajo) de la inicial. ¿Cómo de probable es tener rachas largas? ¿Dónde es más probable que se produzcan, al principio, al final, o por la mitad de la partida? (en otras palabras, ¿cuándo "conviene" retirarse de una partida?). Podemos, por ejemplo, fijar la longitud de las partidas, digamos $N = 2n$, y estudiar cuándo se produce la última "visita al cero" (volver a tener la fortuna inicial). ¿Cuál es la distribución de la máxima fortuna conseguida durante la partida?

7. Dados mágicos

- **Objetivo.** Simular la competición entre varios dados.
- **Modelo.** Lanzamiento de dados "especiales".
- **Esquema de trabajo.** Simulamos el lanzamiento de tres dados: el dado A tiene los números 2, 4, 9; el dado B tiene 3, 5, 7; y C tiene 1, 6, 8. Compiten dos dados: ¿quién gana? ¿Se cumple que si un dado gana a otro, y éste al tercero, entonces el primero gana al tercero (transitividad)?

  Ahora ponemos a competir los tres dados. Podemos también cambiar los pagos del juego: por ejemplo, (a) el premio es 1 para el que gana; (b) el premio es la diferencia de puntos. ¿Y con cuatro dados? ¿Y con n dados?

8. Colas en cajeros

- **Objetivo**. Entender cómo se forman colas en las cajas de un supermercado y tomar decisiones para evitarlo.
- **Modelo.** Simulación aleatoria de las llegadas de los clientes a las cajas y de la atención por parte de los empleados de las cajas.
- **Esquema de trabajo.** Hay tres cajas en un supermercado. Los clientes llegan (en cada unidad de tiempo, un minuto, por ejemplo) a las cajas con una cierta distribución (digamos, una Poisson) y se colocan en ellas aleatoriamente. Los encargados de las cajas atienden m personas (en media) por minuto. ¿Cuál es la longitud de las colas que se forman? ¿Cómo influye si cambiamos la velocidad de atención en una caja? El objetivo es conseguir que, a lo largo de un día, las colas nunca (o al menos no más del 5 % del tiempo) se formen colas mayores que un cierto tope.

  Variación: los clientes, al llegar, se sitúan en la caja cuya cola sea más corta.

9. El viaje ¿sin retorno? del borracho en una, dos y tres dimensiones

- **Objetivo.** Estudiar la probabilidad de que las trayectorias del camino aleatorio en una, dos y tres dimensiones vuelvan al origen.
- **Modelo.** Camino aleatorio en el retículo de 1, 2 y 3 dimensiones.

- **Esquema de trabajo.** Un tipo sale de su casa y se mueve aleatoriamente (con probabilidad 1/2) por los enteros de la recta, a derecha e izquierda. ¿Volverá alguna vez a su casa, al punto de partida? ¿Cuánto tardará, en media?

  Las mismas preguntas si se mueve en el retículo en dos dimensiones, es decir, por los puntos de coordenadas (a, b), donde a y b son enteros (el origen $(0, 0)$ es el punto de partida).

  ¿Qué ocurre si nos movemos en el retículo tridimensional?

  Plantear un modelo (por ejemplo, en dos dimensiones) en el que no nos restrinjamos al retículo.

10. Diseño de ascensores

- **Objetivo.** Tomar una decisión sobre el diseño del funcionamiento de un ascensor en un edificio.
- **Modelo y esquema de trabajo**. Se trata de simular los movimientos de un ascensor, en función de las llamadas que recibe.

  Queremos comparar dos posibles diseños: en uno, el ascensor se queda en el último piso en el que haya dejado pasajeros. En otro, si durante un cierto tiempo no ha sido llamado, baja automáticamente a la planta baja.

  En un modelo sencillo, el edificio tiene 10 pisos, además de la planta baja, y en el ascensor sólo cabe una persona. En un cierto paso, el ascensor está en un determinado piso. Sorteamos primero desde qué piso se le llama (quizás desde ninguno). Y contamos el "tiempo" que tarda en llegar (medido en número de pisos recorridos). Luego sorteamos en qué piso se queda. Y volvemos a empezar.

  El objetivo es minimizar el tiempo medio de espera hasta que llega el ascensor, comparando las posibles estrategias: el ascensor se queda en el último piso visitado, va automáticamente a la planta baja, o quizás hace esto sólo en el caso de no haber sido pedido en el paso anterior.

11. El coleccionista de cromos

- **Objetivo.** Medir el tiempo que se tarda en completar una colección de cromos.
- **Modelo.** Los cromos aparecen con distribución uniforme (e independientemente). Variantes: algunos cromos son especialmente "difíciles".
- **Esquema de trabajo.** La colección consta de n cromos. Cada mañana se visita el kiosko y se compra un sobre que contiene un cromo. Visitamos el kiosko hasta completar la colección. Y medimos el tiempo medio (en número de días) que tardamos en conseguirlo.

12. La ruina de una compañía

- **Objetivo.** Analizar la estrategia que debe seguir una compañía (reservas que debe mantener) para evitar situaciones de quiebra.
- **Modelo.** Los pagos anuales siguen una cierta distribución, por ejemplo, una Poisson (o quizás una lognormal; esto es, el pago anual Y se escribe como $Y = e^X$, donde X es una normal con ciertos parámetros).
- **Esquema de trabajo.** Una compañía de seguros tiene unos ingresos fijos anuales C (las primas que cobra a sus clientes). Los pagos que realiza cada año (por los siniestros sufridos por sus clientes) son, sin embargo, aleatorios. Tiene, además, una reserva de fondos R. El primer año, el balance es $X_1 - C$, donde X_1 son los pagos efectuados ese primer año. Si esa diferencia es mayor que R, entonces la compañía quiebra. Lo mismo ocurre el segundo año, sólo que ahora hay que comparar $X_1 + X_2$ con $2C$. Y así, sucesivamente.

  Marcamos un horizonte temporal, digamos 10 años, y queremos calcular la probabilidad de que la compañía no quiebre en ese periodo, dado un cierto valor de R. O, mejor, determinar el valor de R que hace que la probabilidad de quiebra sea menor de, por ejemplo, un 5 %.

13. Recuentos electorales

- **Objetivo.** Analizar la significación que pueden tener las encuestas parciales (por ejemplo, recuento de las cien primeras papeletas) en el resultado final de unas elecciones.
- **Modelo.** Ballot problem, camino aleatorio.
- **Esquema de trabajo.** En una votación, el candidato A obtiene a votos, mientras que el candidato B obtiene b votos, con $a > b$. Queremos medir la probabilidad de que el candidato finalmente ganador, A, vaya siempre en cabeza a lo largo de todo el recuento (éste es el clásico ballot problem).

  Una variación: sabemos que el candidato A ha conseguido, al final del recuento (hay 1000 votos), un 52 % de los votos, mientras que B ha conseguido un 48 %. ¿Cuál es la probabilidad de que, en el recuento de los primeros 100 votos, el resultado sea el contrario (esto es, que el candidato finalmente ganador tenga menos del 48 % de los votos)?

2. Elección de trabajos. Formación de grupos

Para realizar los trabajos, deberéis formar grupos (de entre 1 y 3 personas). Para poder realizar el trabajo, será necesario haber asistido a un número suficiente de sesiones del Laboratorio. Una vez decidida la composición de cada grupo, se comunicará al profesor de la asignatura, para su aprobación. En esta comunicación constarán los nombres de los integrantes del equipo, así como sus correos electrónicos.

Cada grupo elegirá un trabajo de la lista anterior[1].

Plazo de elección de trabajos y formación de grupos: **del 12 al 19 de diciembre de 2008**.

Dedicaremos las sesiones del Laboratorio del 9 y 16 de enero (las dos últimas del curso) a la realización en el aula de los trabajos.

3. Entrega de trabajos

La fecha límite de entrega de trabajos es el **viernes 23 de enero de 2008**. El trabajo contendrá:

- Una memoria de un par de páginas describiendo el problema en cuestión, el modelo utilizado; y las conclusiones y comentarios sobre los resultados de la simulación.
- Las hojas de cálculo de Excel (o ejecutables de Matlab, o de cualquier otro lenguaje de programación que se desee utilizar) que contengan las simulaciones.
- La memoria se entregará en papel. Los archivos electrónicos se podrán enviar por correo electrónico a `pablo.fernandez@uam.es` o, si fueran excesivamente grandes, se entregarán grabados en un CD.

La nota obtenida en el trabajo se incorporará, en la manera que se dispuso al comienzo del curso, a la nota final de la asignatura de Probabilidad I.

Cabe la posibilidad de que, una vez entregados los trabajos, se convoque a los miembros de algún equipo para que expliquen el trabajo realizado.

Para cualquier duda relacionada con estas instrucciones, o los trabajos en sí, contáctese con `pablo.fernandez@uam.es`.

[1] Si se desea proponer un trabajo que no aparece en la lista, consúltese antes con `pablo.fernandez@uam.es`.