Fundamentos de Matemáticas Financieras

Ignacio Villanueva Díez

Índice general

Introducción

El comienzo de la gran expansión de los mercados de derivados tuvo lugar a principios de los años 70 del siglo XX, no casualmente en el mismo momento en el que Black, Scholes y Merton dieron con un modelo muy verosímil que ajustaba con bastante exactitud los precios de las opciones sobre acciones, el derivado no trivial más usado en esos días (y aún hoy). Dicho modelo fue galardonado más adelante con un premio Nobel, lo que muestra la importancia que la comunidad económica le ha dado. Este modelo, y sus versiones y refinamientos posteriores, utilizan matemática de alto nivel (procesos estocásticos, martingalas, derivadas de Radon-Nikodym, movimiento Browniano, Integración estocástica, Ecuaciones en Derivadas Parciales) muy próxima toda ella al Análisis Matemático, para justificar los precios que el mercado da de ciertos derivados y, lo que es más importante, para estimar con bastante precisión un rango de precios adecuado para un derivado que el mercado no cotiza, o para diseñar estrategias de cobertura para el agente que haya comprado o vendido uno de estos productos. Es decir, no es sólo un buen modelo teórico, sino que las casas de inversión lo usan en su actividad diaria. Además, una vez desarrollados modelos razonablemente fiables de ciertas variables económicas, como la cotización de un activo, la evolución de los tipos de interés, etc., ha sido posible desarrollar metodologías útiles para la medición del *riesgo* asociado a una determinada cartera, y para su cobertura.

El desarrollo simultáneo de los mercados y de los modelos ha dado lugar a la aparición de una gran cantidad de derivados cada vez más sofisticados, explicados a su vez con modelos paralelamente sofisticados. Es a esta especialidad matemática a la que nos referimos con el nombre de *Matemáticas Financieras* o *Finanzas Cuantitativas.*

Capítulo 1

Fundamentos Económicos y planteamiento del problema

Comenzamos este capítulo introductorio exponiendo la noción de derivado financiero, mostrando quiénes son los agentes que actúan en los mercados de derivados y explicando qué ventajas ofrecen para dichos agentes estos mercados frente a los mercados tradicionales de subyacentes. A continuación mostramos la principal noción económica que usaremos en el curso, la noción de *arbitraje*, y mostramos mediante un sencillo ejemplo de valoración de un futuro qué se entiende por *valoración por arbitraje*. Para acabar definimos un derivado más interesante, las opciones, y mostramos cómo se valoran en el modelo más sencillo. Los sucesivos refinamientos de este modelo (y las matemáticas necesarias para cubrirlos) nos ocuparán buena parte del curso.

1. Derivados Financieros

Un *derivado financiero* es un instrumento financiero cuyo valor depende (deriva) del valor de otro (u otros) activo financiero más básico al que se denomina habitualmente *activo subyacente*.

Ilustremos esto con un ejemplo probablemente ya conocido, las *opciones sobre acciones*, (*stock options*): Una opción sobre una acción de, por ejemplo, Telefónica, es un instrumento que te da el derecho (que no la obligación) de comprar un cierto número (supongamos una) de acciones de Telefónica a un cierto precio K en una fecha prefijada T. Ya comentaremos en clase la utilidad de este y otros derivados, baste ahora mencionar que algunos derivados sencillos (sobre todo futuros) se comercializan desde hace mucho tiempo (la tesis de Bachellier, de 1900, estudia entre otros productos algunos derivados que se vendían en la Bolsa de París en su época) pero fue en 1973 cuando el Chicago Board Options Exchange comenzó a comenzó a comercializar opciones sobre acciones de 16 compañías con contratos adecuadamente estandarizados. Desde entonces el mercado de derivados no ha parado de crecer, con la aparición de derivados cada vez más variados y sofisticados, superando en la actualidad en los mercados financieros el volumen de transacciones de derivados a las transacciones realizadas sobre los activos subyacentes.

La pregunta que motiva este curso es la primera que a uno se le ocurre: ¿cuanto cuesta un determinado derivado?. Por supuesto, si el derivado es líquido, cuesta el precio al que se compre y venda en el mercado, pero ¿cuánto *debería* costar?. Supongamos que en este momento una acción de Telefónica cuesta 20 euros. ¿Cuánto debería costar una opción que nos diera derecho a comprar una acción de Telefónica a 25 euros dentro de tres meses?. Veremos a lo largo de este curso que, sorprendentemente, la valoración de esa opción es razonablemente independiente de las expectativas que tengamos sobre el comportamiento futuro de las acciones de Telefónica (comportamiento acerca del cual las previsiones son muy poco fiables) y en cambio depende esencialmente de datos que sí somos capaces de evaluar. Pero las matemáticas necesarias para hacer esta evaluación no son nada triviales. A cambio, no necesitaremos apenas ninguna noción económica, salvo la básica de *arbitraje*.

2. Fundamentos económicos. La noción de Arbitraje

El concepto fundamental que se utilizará para valorar derivados es el de *arbitraje*. Un arbitraje es una oportunidad de ganar dinero sin arriesgar, y se supone que todos los agentes se convertirían en arbitrajistas si tuvieran la oportunidad. Es decir, si en algún momento aparece en el mercado la oportunidad de ganar dinero sin arriesgar, los agentes económicos aprovecharán esa oportunidad: la presión compradora o vendedora de estos agentes hace que los activos involucrados corrijan sus precios hasta que la oportunidad de arbitraje desaparece. Por lo tanto, se supondrá que los mercados trabajan en condiciones de *ausencia de arbitraje* (esta hipótesis a menudo se llama *No Free Lunch*, NFL en la literatura). Esta hipótesis es de hecho bastante realista: las grandes entidades financieras tienen contratados arbitrajistas, personas cuyo única función es vigilar los mercados, detectar arbitrajes y aprovecharlos inmediatamente. Por supuesto su presión compradora o vendedora modifica los precios de manera que la oportunidad de arbitraje rápidamente desaparece, y paradójicamente la misma existencia de los arbitrajistas nos garantiza que podemos suponer Ausencia de Arbitraje en los mercados.

Una breve aclaración. Siempre existe la posibilidad de invertir el dinero en algún instrumento cuya rentabilidad es no aleatoria y conocida de antemano. Ganar dinero por encima de ese tipo de interés es ganar dinero, mientras que obtener una rentabilidad inferior a ese tipo de interés es perder dinero. Por sencillez supondremos a veces que el tipo de interés es 0, pero eso no resta generalidad a los razonamientos.

Los bancos y las grandes entidades financieras consideran que el tipo de interés es el MIBOR, LIBOR, EURIBOR o equivalente. Este índice mide el tipo de interés al que los bancos se prestan dinero entre sí, libre de riesgo.

3. Valoración por arbitraje. Primer ejemplo

Veamos con un ejemplo muy sencillo cómo se utiliza la noción de arbitraje para valorar un derivado financiero. Consideremos un *contrato de futuros* (futures contract) o un *contrato forward*. Un contrato de futuros es un contrato que te obliga a comprar un determinado bien S en una determinada fecha T a un precio K decidido actualmente que es el precio del futuro. El forward contract es esencialmente lo mismo, con pequeñas diferencias formales que ahora mismo no nos interesan. Veamos un ejemplo. Suponed una compañía productora de joyería que utiliza el oro como materia prima. Sabe que el año que viene necesitará 100 Kg de oro, y supongamos (lo que es realista) que la posesión del oro durante un año no le proporciona ninguna renta, mientras que prefiere no tenerlo almacenado durante un año por motivos de seguridad. Esta empresa teme que las fluctuaciones en el precio de oro le puedan causar perdidas económicas, por lo que decide comprar ahora un *futuro* sobre los 100 Kg de oro con vencimiento dentro de un año. Supongamos que el tipo de interés es r anual (es decir, un euro ahora son $(1 + r)$ euros dentro de un año) y que el precio actual del oro es 20.000 euros el Kg. Veamos cuánto debe de costar dicho futuro.

Lo primero que uno piensa es que depende. Recordemos que el contrato de futuros supone la obligación por parte del vendedor de entregar al comprador 100 Kg de oro dentro de un año a cambio de un dinero K, el precio del contrato de futuros, fijado en este momento y que el comprador pagará dentro de un año. Por tanto, uno se puede sentir tentado a pensar que el precio K dependerá de la opinión que cada uno de los agentes tenga acerca del probable precio del oro dentro de un año. Veamos que esto no es así, y que el precio "razonable" de este contrato de futuros se puede determinar con exactitud, independientemente de que el oro vaya a subir o bajar en el próximo año, utilizando tan sólo la noción de arbitraje.

De hecho, el precio del contrato debe de ser $20000 \cdot 100 \cdot (1 + r)$ euros. Veamos por qué. Si el precio del futuro K es superior a esa cifra yo podría pedirle al banco ahora 2,000,000 euros (a un tipo de interés r anual), comprar 100 kg de oro al precio actual, y vender un futuro por los 100 Kg de oro a un precio $K > 2{,}000{,}000(1+r)$. Al cabo del año, el contrato se ejecuta, yo cobro K euros, pago los $2{,}000{,}000(1+r)$ euros al banco y me guardo los $K - 2{,}000{,}000(1+r)$ de beneficio, sin haber aportado ninguna cantidad inicial y sin riesgo por mi parte. Eso es un arbitraje.

Supongamos ahora por el contrario que el precio del futuro K es inferior a $2{,}000{,}000(1+r)$. En ese caso yo compro ahora un futuro por K euros, vendo 100 Kg de oro (para muchos activos, aunque quizás no el oro, uno puede venderlos aunque no los tenga; si esto no os convence, suponed que yo soy un tratante de oro y dispongo de oro que vender) y meto los 2.000.000 de euros que obtengo por la venta en el banco con un interés r libre de riesgo. Al cabo de un año el contrato se cierra, con lo que yo recupero mi oro, saco del banco los $2{,}000{,}000(1+r)$ euros que ahora tengo en él, pago los K euros del contrato y me embolso la diferencia $2{,}000{,}000(1+r) - K$. Eso de nuevo es un arbitraje.

Ejercicio 3.1. *Implementar lo anterior en Excel.*

Por lo tanto, independientemente de la futura cotización del oro, el único precio razonable ahora mismo para ese contrato es $2{,}000{,}000(1+r)$ euros. En este sencillo ejemplo ya se ve la esencia de la *valoración por arbitraje*: Se construye, cuando es posible, una *cartera* de inversión, en este caso formada por ciertas cantidades, positivas o negativas, de oro y dinero en el banco) que *replique* el derivado que estamos intentando valorar. Replicar aquí quiere decir que la cartera se comporta al vencimiento del contrato exactamente igual que el derivado. En ese caso, el valor del derivado es exactamente el valor de la cartera, ya que cualquier otra posibilidad nos lleva a un arbitraje. En nuestro ejemplo la cartera estaba formada por $-2{,}000{,}000$ de euros y $+100$ Kg de oro.

También en este sencillo ejemplo podemos ver que hemos hecho uso implícitamente de varias hipótesis. En primer lugar hemos supuesto que los precios de compra y venta (del oro en este caso) son idénticos. También hemos supuesto que el tipo de interés al que el banco me presta el dinero es el mismo tipo de interés que rige si yo deposito mi dinero en él. Hemos supuesto además que se pueden comprar y vender cantidades ilimitadas de dinero y activos, y que ambos están siempre disponibles, tanto para la compra como para la venta. Todas estas hipótesis son notablemente falsas si uno es un pequeño inversor, pero son bastante ciertas para las grandes entidades financieras e inversoras. De todas formas, en las aplicaciones reales en las que se necesite mayor precisión, se utilizan por separado los precios de compra y de venta (*bid* y *ask*). Se supone también que todos los agentes son "pequeños", en el sentido de que su actuación no supone una modificación apreciable de los precios del mercado.

4. ¿Quiénes son los agentes del mercado?

Para comprender los mercados de derivados, probablemente sea conveniente tener al menos una pequeña idea acerca de los distintos agentes que allí actúan. Los principales son los *hedgers*. Una de las posibles traducciones de *hedge* es cubrir, o cubrirse. Esto es lo que hace el comprador de oro de la sección anterior, cubrirse, o protegerse, ante posibles fluctuaciones en el precio del oro. Por supuesto que eso le puede resultar beneficioso o perjudicial, según que el oro haya subido o bajado en ese año, pero la idea es que el no obtiene sus beneficios especulando con el precio del oro, sino manufacturándolo, por lo que no quiere exponerse a posibles perdidas o ganancias debidas no a su negocio (que el puede más o menos controlar) sino al mercado del oro, sobre el que no tiene ningún control.

Pongamos otro caso, este no exactamente de las finanzas, sino de las técnicas actuariales, aunque ambas están muy relacionadas. Recientemente una aseguradora suiza ha emitido un "bono de catástrofes": para esta aseguradora, una catástrofe es un accidente en el que se ven involucrados 100 o más coches asegurados por ella. En los 10 años anteriores a la emisión del bono sólo habían sucedido 2 catástrofes en Suiza. Obviamente, una de estas catástrofes desequilibra sensiblemente las cuentas de la compañía, por lo que la compañía ha decidido emitir un bono en las siguientes condiciones:

Si en el siguiente año no ocurren catástrofes, el rendimiento del bono es "alto". En cambio si ocurren catástrofes, el rendimiento es menor o incluso negativo. De esta manera para la aseguradora el bono tiene un efecto "amortiguador" de los efectos negativos sobre sus cuentas de una catástrofe.

Hay muchos otros ejemplos de derivados, algunos realmente sofisticados, que a ciertas compañías les interesa comprar o vender, habitualmente con la intención de disminuir su exposición al riesgo de los mercados, y poder así concentrar su esfuerzo en aquello que se supone que la empresa hace bien.

Un carácter totalmente distinto tienen los *especuladores*. Estos son agentes que quieren "apostar" en el mercado. Ellos tienen la opinión de que el precio de ciertos activos va a subir, o bajar, y están dispuestos a arriesgar dinero para respaldar esa opinión, esperando obtener un beneficio. Los especuladores también están interesados en los derivados, porque les permiten “hacer palanca”, en el sentido de que arriesgando la misma cantidad de dinero obtienen beneficios, o perdidas, mucho mayores, que comprando o vendiendo los activos. Veamos un ejemplo.

Supongamos que en este momento un euro cuesta 1’304 dólares y que un especulador piensa que en los próximos tres meses el euro subirá aún más. En ese caso el especulador podría por ejemplo comprar 10.000 euros, (pagando 13.040 dólares), depositar ese dinero en una cuenta bancaria y venderlos al cabo de tres meses.

Otra alternativa posible sería ir mercado de futuros y comprar futuros con vencimiento a tres meses. Supongamos que en este momento el precio de los euros en el mercado de futuros a tres meses es de 1’295 dolares.

Supongamos que en tres meses el euro cuesta 1’320 dolares. En el segundo de los casos, el especulador pagaría 12.950 dolares por los 10.000 euros que en ese momento valdrían 13200 dolares, con lo que ganaría 250 dolares. En el primero de los casos, el especulador vendería los euros por 13.200 dólares y obtendría un beneficio de 160 dolares, que combinado con la renta obtenida por los euros se puede demostrar que le daría un beneficio de 250 dólares (de hecho esa es la forma de valorar los futuros sobre divisas, aunque no lo veremos ahora). Se puede ver también que en el caso de que el euro se hubiera depreciado frente al dolar en esos tres meses las perdidas para el especulador también coinciden. ¿Cuál es entonces la diferencia entre ambas alternativas?. Notad que en el primer caso, el especulador ha requerido una inversión inicial de 13.040 dolares, mientras que en el segundo caso, puesto que el precio del futuro no se paga inicialmente, sólo tiene que hacer una pequeña inversión inicial, (probablemente del orden de los 600 dolares) que utiliza el mercado para garantizar que cumplirá su parte del contrato. Es decir, obtiene el mismo beneficio (o perjuicio) con una inversión inicial 20 veces inferior. En este sentido digo que los derivados le sirven para “hacer palanca”.

EJERCICIO 4.1. *Implementar lo anterior en Excel*

Veamos el caso de las opciones sobre acciones. Como ya hemos visto, una opción de compra sobre una acción es un instrumento que le da al comprador el derecho (no la obligación) de comprar acciones de una determinada compañía a un determinado precio K en un cierto instante T.

Supongamos que es octubre y que una determinada acción de una compañía cuesta 20 euros y que un especulador considera que esta compañía va a revalorizarse en los próximos dos meses. Supongamos que las opciones de compra al precio $K = 25$ dentro de dos meses cuestan 1 euro. Supongamos que el especulador quiere invertir en esta operación 100 euros. Examinemos sus alternativas. Por un lado, puede comprar 5 acciones de la compañia. Por otro, puede comprar 100 opciones de compra. Supongamos que dentro de dos meses las acciones de la compañía cotizan a 35 euros. En ese caso si tenía 5 acciones, puede venderlas y obtener un beneficio de 50 euros. Pero si tenía opciones, puede ejercerlas, comprar 100 acciones a 25 euros y venderlas inmediatamente a 35 euros, con lo que obtiene un beneficio de 1000 euros, diez veces superior al anterior. Por otro lado, si al cabo de dos meses la compañía se cotiza a 15 euros, en el primer caso el especulador habrá perdido 25 euros, mientras que en el segundo caso lo habrá perdido todo (los 100 euros). Es decir, con la misma inversión, en el mercado de derivados se expone a un beneficio o perjuicio mucho mayor que en el mercado de activos.

EJERCICIO 4.2. *Implementar lo anterior en Excel.*

El tercer tipo de agentes que se encuentra en los mercados son los *arbitrajistas.* Estos se dedican a buscar situaciones en las que uno puede, sin arriesgar, obtener un beneficio. Por supuesto esto habitualmente es imposible con los márgenes y comisiones con los que trabaja un pequeño inversor individual, pero las grandes entidades financieras trabajan con tanto comisiones como diferencias entre los precios de compra y venta (*bid-ask spread*) muy pequeñas, lo que permite que eventualmente sean capaces de localizar y aprovechar oportunidades de arbitraje. Como ya dijimos antes, la misma existencia de los arbitrajistas corrige las "imperfecciones" de los mercados, lo que hace que habitualmente se acepte la hipótesis de ausencia de arbitraje como bastante realista.

5. Opciones sobre acciones. El modelo más sencillo

El objetivo de esta sección es presentar el modelo más sencillo (y poco realista) de valoración de opciones, que nos permitirá justificar la necesidad de introducir herramienta matemática más potente para poder refinar el modelo, haciéndolo más realista, y llegar a comprender el método desarrollado por Black, Scholes y Merton en **[2, 3, 11]** para la valoración de opciones de venta sobre acciones (*call stock options*) que hizo a Scholes y Merton (Black ya había muerto) merecedores de un Premio Nobel en 1997.

En todo el capítulo, el mercado consta de un activo sin riesgo (un bono del tesoro, una cuenta bancaria o algo así) con un determinado tipo de interés y de otro activo, las acciones de una compañía, cuyo precio es S unidades monetarias por acción. Suponemos siempre que tenemos crédito ilimitado, que podemos comprar y vender fracciones de acciones, que los precios de compra y venta coinciden y que el interés al que el banco presta dinero es el mismo interés que paga por tener dinero depositado en él. También se supone que no hay comisiones de compra y que uno puede vender acciones aunque no las tenga. Todas estas suposiciones son muy poco realistas para un pequeño inversor, pero bastante realistas para grandes inversores. Respecto al tipo de interés, si os molesta en los cálculos, siempre se puede suponer cero sin perder generalidad, y os aconsejo que así lo hagáis en una primera lectura.

También todo el capítulo sucederá en "tiempo discreto", en contraposición al "tiempo continuo" cuyo estudio requiere de herramienta matemática más sofisticada que estudiaremos más adelante.

6. El modelo más simple

Suponemos inicialmente que sólo consideramos dos instantes, $T = 0$ y $T = 1$. En $T = 0$ nuestro activo (las acciones) valen S. Suponemos que en $T = 1$ las acciones sólo han podido hacer dos cosas: o bien han subido, en cuyo caso valdrán exactamente S_u, o bien han bajado, y en ese caso valdrán S_d. Suponemos también que el tipo de interés para el periodo es r, es decir, 1 euro depositado en el banco en $T = 0$ valdrá $1 + r$ euros en $T = 1$.

Nuestro problema es encontrar cuál debe de ser el valor de una opción de compra (call stock option) que da el derecho a comprar en $T = 1$ una acción a un precio K.

Veamos qué puede pasar. Llamando S_1 al precio de la acción en $T = 1$, el valor de la opción en $T = 1$ es

$$(S_1 - K)^+ := \max\{S_1 - K, 0\};$$

en efecto, si $S_1 - K > 0$, en $T = 1$ ejercemos el derecho de compra que nos da la opción y compramos una acción al precio K que inmediatamente vendemos al precio de mercado S_1, con lo que nos embolsamos un beneficio $S_1 - K$ (en la práctica lo que ocurre es que el vendedor de la acción le paga al comprador esa cantidad directamente). En cambio, si $K \geq S_1$ no nos interesará ejercer el derecho de compra (ya que podemos encontrar las acciones en el mercado a un precio más barato) por lo que la opción no tendrá ningún valor.

Aunque la observación que sigue no es estrictamente necesaria para los razonamientos que vienen a continuación, observemos que en este ejemplo que las únicas condiciones "razonables" para el modelo son

$$S_d \leq S(1+r) \leq S_u$$

y

$$S_d \leq K \leq S_u.$$

Para justificar las primeras desigualdades, nótese que si S_d fuera mayor que $S(1+r)$ comprar acciones proporcionaría un arbitraje (un beneficio seguro superior al que da el banco) y si S_u fuera menor que $S(1+r)$ vender acciones también proporcionaría un arbitraje; ya hemos comentado que supondremos siempre que los mercados funcionan en *ausencia de arbitraje.*

Para justificar las segundas desigualdades, nótese que si $K < S_d$, entonces la opción siempre se ejercería, mientras que si $K > S_u$ la opción nunca se ejercería. En este segundo caso el único precio razonable para la opción es 0.

Volvamos a nuestro problema. ¿Cuál debe de ser el valor de la opción en $T = 0$?. La idea "buena" es *replicar* la opción con una *cartera de cobertura (hedging portfolio)* compuesta por distintas cantidades (positivas o negativas) de los activos del mercado (en este caso tan sólo dinero en el banco y acciones) *cuyo precio en* $T = 0$ *es conocido* y cuyo comportamiento en $T = 1$ sea exactamente el mismo que el de la opción. Si logramos replicar así la opción, es claro que el precio de la opción en $T = 0$ deberá coincidir con el precio del hedging portfolio.

Veamos cómo podemos hacer esto en nuestro sencillo modelo inicial.

Una cartera en este caso es un par (α, β) formada por α euros en el banco y β acciones. Ambos pueden ser positivos o negativos. Veamos como se porta nuestra cartera en $T = 1$. Si las acciones han subido, nuestra cartera valdrá

$$(1+r)\alpha + \beta S_u$$

mientras que si las acciones han bajado nuestra cartera valdrá

$$(1+r)\alpha + \beta S_d.$$

Definamos

$$C_u := \text{máx}\{0, S_u - K\}$$

y

$$C_d := \text{máx}\{0, S_d - K\},$$

como lo que pasa a valer la opción en $T = 1$ en función de lo que haya ocurrido con la acción. Observad que en las condiciones "razonables" mencionadas anteriormente tenemos $C_u = S_u - K$ y $C_d = 0$.

Si queremos que nuestra cartera se porte en $T = 1$ exactamente como la opción, se ha de cumplir

$$\begin{cases} (1+r)\alpha + \beta S_u = C_u \\ \\ (1+r)\alpha + \beta S_d = C_d \end{cases}$$

Resolviendo el sistema es fácil ver que la única solución es

$$\alpha^* = \frac{1}{1+r}\left(C_d + \frac{S_d(C_d - C_u)}{S_u - S_d}\right), \quad \beta^* = \frac{C_u - C_d}{S_u - S_d}$$

Y por tanto, el hedging portfolio vale en $T = 0$

$$q = \alpha^* + \beta^* S = \frac{1}{1+r}\left(pC_u + (1-p)C_d\right),$$

donde

$$p := \frac{1}{S_u - S_d}\left((1+r)S - S_d\right).$$

Claramente, puesto que el hedging portfolio se porta en $T = 1$ exactamente igual que la opción, ambos deberían valer lo mismo en $T = 0$, por lo que el único precio posible para la opción es q. Aunque esto es totalmente obvio, veámoslo usando razonamientos de arbitraje:

Si la opción valiera $q' > q$, venderíamos una opción. Con q de los q' euros obtenidos por la venta de la acción construimos una cartera (α^*, β^*). Los restantes $q' - q$ euros serán nuestra beneficio, ya que en $T = 1$ nuestra cartera se va a portar exactamente como la opción (independientemente de que la acción haya subido o bajado) y por tanto nos cubren (hedge) frente a los costes a los que tendremos que hacer frente como vendedores de la opción.

Si por el contrario la opción valiera $q' < q$ construiríamos la cartera $(-\alpha^*, -\beta^*)$ (si por ejemplo $\alpha^*{>}0$ y $\beta^* > 0$, la cartera $(-\alpha^*, -\beta^*)$ consiste en pedirle prestado al banco α euros y vender β acciones) lo que nos proporcionaría q euros con los que podríamos comprar una opción y guardarnos un beneficio $q - q'$. En $T = 1$, tanto si la acción ha subido como si ha bajado, el resultado de la opción nos permite deshacer nuestra cartera, y terminamos con el beneficio $q - q'$ (multiplicado por $1 + r$).

La siguiente observación es importante. Hemos dicho que la opción vale

$$q = \frac{1}{1+r}\left(pC_u + (1-p)C_d\right),$$

donde

$$p := \frac{1}{S_u - S_d}\left((1+r)S - S_d\right).$$

Puesto que siempre $0 \leq p \leq 1$, podemos definir una probabilidad P sobre "el conjunto de los estados del mundo en $T = 1$" de manera que p es "la probabilidad P de que la acción suba" y $1 - p$ es "la probabilidad de que la acción baje", con esa interpretación, y suponiendo momentáneamente $r = 0$, se tiene que el valor de la opción es la esperanza matemática de su valor futuro con respecto a la probabilidad P, o dicho de otra forma,

$$q = E_P[C] = \int C dP$$

donde C es el valor de la opción "en función de los estados del mundo". Quizás estos comentarios no son totalmente formales, pero sí son totalmente formalizables y encierran la dirección en la que nos debemos mover para resolver nuestro problema en modelos más complejos.

La forma más sencilla de introducir el tipo de interés en nuestra interpretación anterior es simplemente acostumbrarse a que si el tipo de interés es r, entonces 1 euro de $T = 0$ es "lo mismo" que $1 + r$ euros de $T = 1$, de forma que dividir por $1 + r$ un cierto número de euros de $T = 1$ no es más que hallar el *valor descontado* de esos euros en $T = 0$.

Capítulo 2

Procesos Estocásticos en tiempo discreto

En este capítulo, de carácter técnico, introducimos la teoría de procesos estocásticos en tiempo discreto. Estudiamos qué son procesos estocásticos discretos, qué entendemos por una *martingala* (para lo cual hemos definido anteriormente la *esperanza condicional*) y finalmente estudiamos las propiedades básicas de las martingalas que vamos a necesitar más adelante.

1. Notación

Recordamos a continuación un poco de notación y definiciones básicas de Teoría de la Medida.

Consideraremos habitualmente un espacio Ω. Dentro del conjunto de partes de Ω consideramos una o varias *σ-álgebras* ($\mathcal{F}$, $\mathcal{F}_i$). Una *probabilidad* definida sobre Σ es una función

$$P : \mathcal{F} \longrightarrow [0,1]$$

contablemente aditiva.

Dados $(\Omega, \mathcal{F})$, decimos que

$$f : \Omega \longrightarrow \mathbb{R}$$

es una *función medible* (para los analistas) o una *variable aleatoria* (para los probabilistas) si $f^{-1}(B) \in \mathcal{F}$ para todo $B \in \mathcal{B}(\mathbb{R})$, la σ-álgebra de los conjuntos de Borel de $\mathbb{R}$. En esa definición se puede sustituir $\mathbb{R}$ por cualquier espacio medible, en particular por $\mathbb{R}^n$ con la σ-álgebra de Borel.

Dos funciones medibles f, g que coincidan salvo en un conjunto de de medida (probabilidad) nula se dicen iguales en casi todo punto (c.t.p.). Puesto que la probabilidad no va a ser capaz de distinguir f de g, lo habitual es identificarlas (es decir, cocientar el espacio de funciones medibles en que estemos por la relación de equivalencia "coincidir c.t.p.").

En particular, esto se hace con los espacios (que supongo ya conocidos) $\mathcal{L}_p(\Omega, \mathcal{F}, P)$, que pasamos a llamar $L_p(\Omega, \mathcal{F}, P)$ una vez pasado al cociente.

2. La esperanza condicional

El operador *esperanza condicional* va a ser de gran importancia para nosotros. La manera habitual de presentarlo en los cursos de Teoría de la Medida, influidos por el Análisis Funcional, es por medio del Teorema de Radon-Nikodym. En esta breve sección lo presentaré de una forma que considero más elemental, siguiendo [**10**], donde se pueden consultar más detalles.

Supongamos fijado un espacio de probabilidad $(\Omega, \mathcal{F}, P)$, y consideremos el espacio $L_p(\Omega, \mathcal{F}, P)$, que abreviaré en lo que sigue como $L_p(\mathcal{F})$ o simplemente $\mathcal{L}_p$. Sea $\mathcal{G} \subset \mathcal{F}$ una sub-σ-álgebra. En ese caso, $P_{|\mathcal{G}}$ es una probabilidad y podemos considerar el espacio $L_p(\mathcal{G}) := L_p(\Omega, \mathcal{G}, P_{|\mathcal{G}})$. Es fácil ver que $L_p(\mathcal{G}) = L_p(\mathcal{F}) \cap L_0(\mathcal{G})$.

También es fácil ver que $L_p(\mathcal{G})$ es un subespacio cerrado para la norma, y que es un subretículo cerrado para la convergencia monótona (esto es, si $(f_n)_n \subset L_p(\mathcal{G})$ y $(f_n) \nearrow f$ entonces $f \in L_p(\mathcal{G})$) y que contiene a la función $\mathbb{1}$. Recíprocamente, estas dos últimas propiedades caracterizan los subespacios de L_p de la forma $L_p(\mathcal{G})$: En efecto, sea $M \subset L_p$ un subespacio cerrado, que simultáneamente es un subretículo cerrado para la convergencia monótona y que contiene a la función $\mathbb{1}$. Sea

$$\mathcal{G} := \{G \in \mathcal{F} \text{ tales que } \chi_G \in M\}$$

No es difícil ver que $\mathcal{G}$ así definida es una σ-álgebra, y es claro que $L_p(\mathcal{G}) \subset M$. Por otro lado, si $f \in M$, entonces también

$$\mathbb{1}_{f>0} := \lim_n (nf^+ \wedge \mathbb{1}) \in M$$

por las propiedades de M. Por lo tanto el conjunto

$$\{f > 0\} := \{\omega : f(\omega) > 0\} \in \mathcal{G}$$

Puesto que $\mathbb{1} \in M$, para todo $\lambda \in \mathbb{R}$ también

$$\{f > \lambda\} = \{f - \lambda\mathbb{1} > 0\} \in \mathcal{G}.$$

A partir de ahí, los argumentos estándar de aproximación de funciones en L_p por funciones simples nos permiten concluir que $f \in L_p(\mathcal{G})$, y por tanto que $M \subset L_p(\mathcal{G})$ y $M = L_p(\mathcal{G})$.

En cuanto comencemos a estudiar procesos estocásticos resultará aparente la importancia de los subespacios $L_p(\mathcal{G})$ de L_p. La idea es más o menos la siguiente: una función $f \in L_p$ es lo que en física, particularmente en cuántica, se suele denominar un *observable*, es decir, algo que podemos "observar" para determinar el estado de nuestro sistema. Habitualmente es razonable pedirle al conjunto de los observables que sean un subespacio vectorial cerrado, a la vez que un subretículo cerrado para la convergencia monótona, así como que contenga a las funciones constantes (los múltiplos de $\mathbb{1}$). Por tanto, los razonamientos anteriores nos dicen que en esos casos los observables forman precisamente un subespacio de la forma $L_p(\mathcal{G})$. En el caso que a nosotros nos interesará, la σ-álgebra $\mathcal{F}$ representará "lo que se sabe, o se puede saber" en cierto instante final, mientras que $\mathcal{G}$ será lo que se sabe en un cierto instante anterior T. El hecho de que la información se preserva (hay memoria) se refleja en que $\mathcal{G} \subset \mathcal{F}$. Un observable en T es alguna función $L_p(\mathcal{G})$.

Parece más o menos claro que resultaría interesante disponer de "operadores de predicción"

$$U : L_p(\mathcal{F}) \longrightarrow L_p(\mathcal{G})$$

que nos permitan estudiar $L_p(\mathcal{F})$ por medio de su imagen en $L_p(\mathcal{G})$.

Estos operadores U deben preservar la función $\mathbb{1}$ (lo que podemos observar de una función constante es así mismo constante), deben preservar la estructura de retículo (lo que podemos observar de una función positiva debe de ser positivo) y no debe de aumentar la norma.

Es fácil ver que si un operador $U : L_p \longrightarrow L_p$ verifica las condiciones anteriores entonces su conjunto de puntos fijos es un subespacio M cerrado para la norma, a la vez que un subretículo cerrado para la convergencia monótona que contiene a la función $\mathbb{1}$. Por tanto, existe una σ-álgebra $\mathcal{G}$ de manera que U es precisamente una proyección sobre $L_p(\mathcal{G})$.

De esa forma, la "predicción" dada por U es precisamente la información obtenida al "observar" $L_p(\mathcal{G})$.

Con esta introducción como motivación, ya podemos introducir la *esperanza condicional.*

Supongamos que estamos en el caso $p = 2$ (para tener a nuestra disposición la proyección ortogonal). Definimos entonces el operador esperanza condicional

$$\mathbf{E}(\cdot|\mathcal{G}) : L_2(\mathcal{F}) \longrightarrow L_2(\mathcal{G})$$

como la proyección ortogonal. Por tanto, por las propiedades de la proyección ortogonal, para todos $f \in L_2(\mathcal{F})$, $g \in L_2(\mathcal{G})$,

$$\langle \mathbf{E}(f|\mathcal{G}), g \rangle = \int_\Omega \mathbf{E}(f|\mathcal{G}) g dP = \int_\Omega f g dP = \langle f, g \rangle.$$

Esta igualdad además caracteriza a la esperanza condicional, y es claro que es suficiente comprobarla con las funciones características, de manera que $\mathbf{E}(f|\mathcal{G})$ es la única función de $L_2(\mathcal{G})$ que verifica que, para

todo $G \in \mathcal{G}$
$$\int_G \mathbf{E}(f|\mathcal{G})dP = \int_G f dP$$

Es un ejercicio comprobar que la esperanza condicional cumple lo que esperábamos que cumplieran los "operadores de predicción " U mencionados anteriormente.

A menudo en matemática financiera podremos elegir la norma p con la que trabajamos, de manera que es frecuente elegir $p = 2$ para trabajar en un espacio de Hilbert. En cualquier caso, la esperanza condicional así definida se puede extender a todo $L_1(\mathcal{F})$ "a mano" (sin necesidad de utilizar el Teorema de Radon-Nikodym). La idea de la extensión se puede ver en [**10**], y con todo detalle por ejemplo en [**12**].

Resumimos en el siguiente resultado las principales propiedades de la esperanza condicional que vamos a utilizar.

TEOREMA 2.1. *Sean $\Omega, \mathcal{F}, \mathcal{G}, P$ como en los razonamientos anteriores. Para todo $f \in L_1(\mathcal{F})$ existe una única $\boldsymbol{E}(f|\mathcal{G}) \in L_1(\mathcal{G})$ tal que, para todo $G \in \mathcal{G}$,*
$$\int_G \boldsymbol{E}(f|\mathcal{G})dP = \int_G f dP.$$

El operador $\boldsymbol{E}(\cdot|\mathcal{G}) : L_1(\mathcal{F}) \longrightarrow L_1(\mathcal{G})$ se llama esperanza condicional *y tiene las siguientes propiedades:*

1. *$\boldsymbol{E}(\cdot|\mathcal{G})$ es lineal, positivo, idempotente, de norma 1 y $\boldsymbol{E}(\mathbb{1}|\mathcal{G}) = \mathbb{1}$.*
2. *(Convergencia monótona) Si $(f_n) \subset L_1(\mathcal{F})$, $f \in L_1(\mathcal{F})$ y $f_n \nearrow f$ en L_1, entonces*
$$\boldsymbol{E}(f_n|\mathcal{G}) \nearrow \boldsymbol{E}(f|\mathcal{G})$$
en $L_1(\mathcal{G})$.
3. *(Fatou) Si $(f_n) \subset L_1(\mathcal{F})$, $f, g \in L_1(\mathcal{F})$ y $g \leq f_n$ para todo $n \in \mathbb{N}$, entonces*
$$\boldsymbol{E}(\liminf f_n|\mathcal{G}) \leq \liminf \boldsymbol{E}(f_n|\mathcal{G});$$
Si $f_n \leq g$ para todo $n \in \mathbb{N}$, entonces
$$\boldsymbol{E}(\limsup f_n|\mathcal{G})| \geq \limsup \boldsymbol{E}(f_n|\mathcal{G})$$
4. *(Valor medio) Si $h \in L_\infty(\mathcal{G})$, entonces*
$$\boldsymbol{E}(fh|\mathcal{G}) = h\boldsymbol{E}(f|\mathcal{G}).$$
5. *(Tower property) Si $\mathcal{H} \subset \mathcal{G}$ es una sub-σ-álgebra, entonces*
$$\boldsymbol{E}(\boldsymbol{E}(f|\mathcal{G})|\mathcal{H}) = \boldsymbol{E}(f|\mathcal{H}) = \boldsymbol{E}(\boldsymbol{E}(f|\mathcal{H})|\mathcal{G})$$

Demostración. Suponemos cierto el apartado 1, que consiste en extender la esperanza condicional a todo L_1.

Para ver el apartado 2, sea $G \in \mathcal{G}$, y $(f_n), f$ como en la hipótesis. Entonces, el Teorema de la Convergencia Dominada nos dice que

$$\int_G f_n dP \to \int_G f dP,$$

y de ahí se sigue el resultado.

Análogamente se demuestra el apartado 3.

Veamos el apartado 4 primero en el caso del operador esperanza condicional de $L_2(\mathcal{F})$ en $L_2(\mathcal{G})$. Es decir, suponemos $f \in L_2(\mathcal{F})$. Observamos primeramente que $g\mathbf{E}(f|\mathcal{G}) \in L_2(\mathcal{G})$.

Para todo $h \in L_2(\mathcal{G})$ se tiene que $gh \in L_2(\mathcal{G})$ y, por la definición

$$\int hg\mathbf{E}(f|\mathcal{G})dP = \int hgf dP,$$

es decir

$$\langle h, g\mathbf{E}(f|\mathcal{G}) - gf\rangle = 0.$$

Es decir, $g\mathbf{E}(f|\mathcal{G}) - gf$ es ortogonal a $L_2(\mathcal{G})$, lo que nos permite acabar. Una vez probado 4 en L_2, las mismas técnicas que nos permitieron extender la definición de esperanza condicional a todo L_1 nos permiten extender 4.

El apartado 5 se demuestra con facilidad, y visto en el caso particular de L_2 tiene una interpretación geométrica fácil y clara. □

3. Procesos Estocásticos Discretos

Suponemos fijada una *base estocástica* $(\Omega, \Sigma, P, (\Sigma_n), \mathbb{N})$, que consiste en un espacio de probabilidad (Ω, Σ, P) junto con una *filtración* (filtration) (Σ_n) indexada en el caso discreto por $\mathbb{N}$. Nosotros trabajaremos habitualmente en el caso discreto y *finito*, de manera que el conjunto de índices será a menudo $\{0, 1, 2, \ldots, N\}$. En cualquier caso, la filtración (Σ_n) consiste en una colección creciente de σ-álgebras

$$\Sigma_0 \subset \Sigma_1 \subset \cdots$$

donde se suele suponer que Σ_0 contiene a todos los conjuntos de medida nula y que Σ es la σ-álgebra generada por $\cup_n \Sigma_n$. Las sucesivas σ-álgebras representan sucesivos refinamientos de nuestro conocimiento. En nuestro caso, este refinamiento del conocimiento se irá dando con el paso del tiempo, de manera que las sucesivas σ-álgebras representarán "lo que se sabe" en los sucesivos instantes temporales t_n.

Una vez fijada una base estocástica $(\Omega, \Sigma, P, (\Sigma_n), \mathbb{N})$ definimos un *proceso estocástico* como una sucesión (en el caso discreto) $X = (X_n)$ de variables aleatorias (es decir, funciones Σ-medibles). Nuestros procesos estocásticos tomarán habitualmente valores en $\mathbb{R}$ o en $\mathbb{R}^n$.

Un proceso estocástico $X = (X_n)$ se dice *adaptado* a la filtración (Σ_n) si para todo n, X_n es Σ_n-medible.

Un proceso estocástico (X_n) se dirá *predecible* si para todo $n \geq 1$ X_n es Σ_{n-1}-medible, y X_0 es Σ_0-medible.

Un comentario quizás sea conveniente. A menudo lo que uno conoce es precisamente el proceso estocástico, y no la filtración. Si nuestro conocimiento del mundo viene a través de uno o varios procesos estocásticos, tiene sentido definir la filtración como las mínimas σ-álgebras que hacen que el proceso sea adaptado.

Aunque lo veremos con más detalle más adelante, comentaré brevemente algunos de los procesos estocásticos con que nos vamos a encontrar de forma natural en nuestros modelos.

Dado un activo cualquiera (una acción, oro, dólares, etc) modelizaremos su cotización en función del tiempo (y de lo que pase en el mundo, es decir, de $\omega \in \Omega$) como un proceso estocástico $S = (S_n)$ adaptado a la filtración Σ_n que representa "lo que ha sucedido" o "lo que se sabe" hasta el instante t_n. Claramente, S debe de ser adaptado a Σ_n, ya que la cotización sólo puede depender de lo que ya ha ocurrido (nadie tiene bolas de cristal).

Dado un agente, podrá diseñar una cartera θ. Típicamente, esta cartera no será constante en el tiempo, sino que nuestro agente la irá modificando en función de lo que ya haya sucedido. Veremos que resultará natural pedirle a la cartera que sea un proceso estocástico $\theta = (\theta_n)$ *predecible*.

También es posible pensar en un proceso estocástico discreto con valores reales como una función

$$X : \Omega \times \mathbb{N} \longrightarrow \mathbb{R}$$

donde

$$X(\omega, n) = X_n(\omega)$$

Si fijamos $n_0 \in \mathbb{N}$, obtenemos la variable aleatoria X_{n_0}. Si en cambio fijamos $\omega_0 \in \Omega$ la función

$$X(\omega_0) : \mathbb{N} \longrightarrow \mathbb{R}$$

que obtenemos se suele llamar la trayectoria o el camino (path) de ω_0, aunque esta notación quizás se utiliza más para procesos estocásticos en tiempo *continuo*.

4. Martingalas en tiempo discreto

Vamos a concentrarnos momentáneamente en el estudio de un tipo especial de procesos estocásticos de gran importancia para nosotros, las martingalas.

Definición 4.1. Dada una base estocástica como antes, un proceso estocástico adaptado $M = (M_n)$ se dice una *martingala* si $M_n \in L_1(P)$ para todo n y además, para todo $0 \leq n \leq m$ se tiene

$$\mathbf{E}(M_m|\Sigma_n) = M_n.$$

X se dice una *supermartingala* si $\mathbf{E}(M_m|\Sigma_n) \leq M_n$ y una *submartingala* si $\mathbf{E}(M_m|\Sigma_n) \geq M_n$.

Un par de observaciones:

En primer lugar, notemos que, al contrario que en la definición de proceso estocástico, en la definición de martingala la probabilidad P elegida *sí* es relevante, ya que es muy fácil dar ejemplos (y los veremos de forma natural en nuestros modelos financieros) de procesos que son una martingala respecto de cierta medida, y no lo son respecto de otra.

Notemos también que utilizando la Tower Property de la esperanza condicional, queda claro que M es una martingala si y sólo si, para todo n,

$$M_n = \mathbf{E}(M_{n+1}|\Sigma_n).$$

Es importante observar también que un proceso estocástico *predecible* es una martingala si y sólo si es constante. En efecto, puesto que M_{n+1} es Σ_n medible, usando la propiedad 4 del Teorema 2.1 y la definición de martingala, se tiene que, para todo n,

$$M_{n+1} = M_{n+1}\mathbf{E}(\mathbb{1}|\Sigma_n) = \mathbf{E}(M_{n+1}|\Sigma_n) = M_n.$$

Para empezar a interpretar la noción de martingala, pensemos en el siguiente ejemplo, que más adelante veremos que tiene sentido económico. Supongamos que $S = (S_n)$ es un proceso estocástico que modeliza la cotización de una determinada acción, de forma que S_n es la variable aleatoria que nos indica la cotización (no determinista) de la acción en el instante n. Supongamos además que la σ-álgebra Σ_n nos muestra lo que es sabido en el instante n. Supongamos que existe una medida P respecto de la cual S es una martingala. Supongamos que estamos en el instante n y queremos hacer una predicción acerca de lo que ocurrirá en el instante $m > n$. Esta predicción es

$$\mathbf{E}(S_m|\Sigma_n)$$

y por ser S una martingala con respecto de P, se tiene

$$\mathbf{E}(S_m|\Sigma_n) = S_n,$$

es decir, la mejor predicción que podemos hacer sobre los precios es que seguirán como están hoy.

Veremos que lo que el Teorema Fundamental de Valoración de Activos nos dice es precisamente que existe una medida respecto de la cual el proceso de precios es una martingala si y sólo si el mercado está libre de arbitraje.

El ejemplo más claro de martingala es probablemente el más trivial, aunque veremos que contiene en cierto sentido la idea principal de las martingalas.

EJEMPLO 4.2. *Sea $f \in L_1$ una variable aleatoria Σ-medible. Sea $M_n := \boldsymbol{E}(f|\Sigma_n)$. Entonces el proceso $M = (M_n)$ es una martingala respecto de P.*

Resulta un ejercicio sencillo (trivial) demostrar que M así definido es una martingala. De hecho, se puede probar que si las variables aleatorias (M_n) son *uniformemente integrables*, entonces toda martingala es de esta forma [**10**]. No incluimos aquí el enunciado de familia uniformemente integrable de funciones, simplemente recalcaré que en el caso en que nosotros vamos a trabajar inicialmente (Ω finito y tiempo discreto y finito) toda martingala es uniformemente integrable, luego toda martingala en el fondo no es sino las sucesivas proyecciones de una única variable aleatoria. Esto justifica que, cuando uno trabaja con procesos estocásticos, uno piense a menudo en las martingalas como el equivalente a las constantes.

Desde el punto de vista de la modelización de juegos de azar, las martingalas representan los juegos "justos", o quizás debiéramos decir más bien "indiferentes". Si $M = (M_n)$ representan las ganancias acumuladas del jugador tras n jugadas, el hecho de que

$$\mathbf{E}(M_m|\Sigma_n) = M_n$$

se interpreta como que la mejor predicción que podemos hacer desde el instante n de las futuras ganancias del jugador en el instante m es que en dicho instante tendrá exactamente las mismas ganancias que en n, es decir, resulta indiferente (en promedio) que siga jugando o que no juegue más. Bajo esta interpretación, las supermartingalas representan juegos desfavorables al jugador, y las submartingalas juegos favorables.

5. Transformadas por martingalas (martingale transforms)

Las martingale transforms son el análogo discreto de las integrales estocásticas. Si comprendemos el papel de las transformadas en el modelo discreto estaremos en mejores condiciones para entender mejor las integrales estocásticas, que presentan una dificultad matemática sensiblemente superior.

DEFINICIÓN 5.1. Sea $\theta = (\theta_n)$ un proceso *predecible* y $M = (M_n)$ un proceso *adaptado*. La *transformada de M por θ* es el proceso $G = \theta \bullet M$ dado por

$$G_n = \sum_{i=1}^{n} \theta_i (M_i - M_{i-1}) = \sum_{i=1}^{n} \theta_i \Delta M_i$$

OBSERVACIÓN 5.2. *También se puede definir la transformada como*

$$G_n \theta_0 M_0 + \sum_{i=1}^{n} \theta_i (M_i - M_{i-1})$$

pero para nosotros será más cómoda la primera definición.

La propiedad más interesante para nosotros de las transformadas es que mantienen las martingalas. Antes de demostrar ese resultado enunciamos y probamos un sencillo lema auxiliar.

LEMA 5.3. *Un proceso estocástico M es una martingala si y sólo si, para todo $n > 0$*

$$\boldsymbol{E}(\Delta M_n | \Sigma_{n-1}) = 0$$

DEMOSTRACIÓN. En efecto,

$$\mathbf{E}(\Delta M_n | \Sigma_{n-1}) = \mathbf{E}(M_n - M_{n-1} | \Sigma_{n-1}) =$$
$$\mathbf{E}(M_n | \Sigma_{n-1}) - \mathbf{E}(M_{n-1} | \Sigma_{n-1}) = \mathbf{E}(M_n | \Sigma_{n-1}) - M_{n-1}$$

□

TEOREMA 5.4. *Si M es una martingala y θ es un proceso predecible tales que $(\theta \bullet M)_n \in L_1$ para todo n, entonces la transformada $G = \theta \bullet M$ es una martingala.*

DEMOSTRACIÓN. Usando el Lema 5.3, basta observar que, puesto que θ es predecible

$$\mathbf{E}(\Delta(\theta \bullet M)_n | \Sigma_{n-1}) = \mathbf{E}(\theta_n (M_n - M_{n-1}) | \Sigma_{n-1}) = \theta_n \mathbf{E}(\Delta M_n | \Sigma_{n-1}) = 0.$$

□

Da ahí se sigue la siguiente caracterización, sencilla pero útil, de las martingalas.

También ahora necesitamos un sencillo lema previo.

LEMA 5.5. *Dado un espacio de probabilidad* (Ω, Σ, P) *y una* σ-*álgebra* $\Sigma' \subset \Sigma$, *para toda* $f \in L_1(\Sigma)$ *se tiene*

$$\boldsymbol{E}(\boldsymbol{E}(f|\Sigma')) = \boldsymbol{E}(f).$$

DEMOSTRACIÓN. En efecto, utilizando el Teorema 2.1, y observando que siempre $\Omega \in \Sigma'$, tenemos que

$$\mathbf{E}(\mathbf{E}(f|\Sigma')) = \int_\Omega \mathbf{E}(f|\Sigma')dP = \int_\Omega f dP = \mathbf{E}(f).$$

□

PROPOSICIÓN 5.6. *Un proceso adaptado* M *es una martingala si y solo si para todo proceso predecible y acotado* θ *se tiene que, para todo* $n \geq 1$,

$$\boldsymbol{E}((\theta \bullet M)_n) = \boldsymbol{E}\left(\sum_{j=1}^{n} \theta_j \Delta M_j\right) = 0.$$

En el caso de que el conjunto de índices sea discreto y finito $0, 1, \ldots, T$, *para que* M *sea una martingala es suficiente con que* $\boldsymbol{E}((\theta \bullet M)_T) = 0$

DEMOSTRACIÓN. Si M es una martingala, también lo es su transformada $\theta \bullet M$.

Entonces, utilizando el lema anterior, para todo $1 \leq j \leq n$ se tiene

$$\mathbf{E}(\theta_j \Delta M_j) = \mathbf{E}(\theta_j M_j - \theta_j M_{j-1}) = \mathbf{E}(\mathbf{E}(\theta_j M_j - \theta_j M_{j-1}|\Sigma_{j-1})) =$$
$$= \mathbf{E}(\mathbf{E}(\theta_j M_j|\Sigma_{j-1})) - \mathbf{E}(\mathbf{E}(\theta_j M_{j-1}|\Sigma_{j-1})) = \mathbf{E}(\theta_j M_{j-1} - \theta_j M_{j-1}) =$$
$$= \mathbf{E}(0) = 0$$

Recíprocamente, si la propiedad es cierta para todo θ, sea $m > 0$, sea $A \in \Sigma_{m-1}$ y sea θ el proceso predecible dado por

$$\theta_n = \begin{cases} 0 & \text{si } n \neq m \\ \chi_A & \text{si } m = n \end{cases}$$

Entonces, para todo $n > m$ se tiene

$$0 = \mathbf{E}((\theta \bullet M)_n) = \mathbf{E}(\chi_A(M_m - M_{m-1})).$$

Puesto que esto es cierto para todo $A \in \Sigma_{m-1}$, se sigue que

$$\mathbf{E}(\Delta M_m|\Sigma_{m-1}) = 0,$$

por lo que M es una martingala.

Se ve fácilmente que la demostración anterior demuestra también la última afirmación de la proposición. □

No diremos más ahora de la Teoría de Martingalas, puesto que esto resulta suficiente para nuestro primer acercamiento, razonablemente elemental. Para leer más acerca de esta Teoría se puede consultar por ejemplo [**10**], y las referencias que allí se mencionan.

Capítulo 3

Opciones sobre acciones. Un modelo en tiempo discreto

En este capítulo utilizaremos la herramienta desarrollada en el Capítulo 2 para construir un modelo bastante general y realista de los mercados de derivados cuando se supone que el tiempo es discreto.

1. La estructura de la información

Trabajaremos en tiempo discreto y horizonte finito. Es decir, suponemos un conjunto de instantes

$$\mathbb{T} = \{0, 1, 2, \ldots, T\}$$

de manera que sólo en esos instantes "ocurre algo", es decir, son los únicos momentos en los que cambian los precios, se puede comprar o vender, etc.

Suponemos también la existencia de un espacio de medida (Ω, Σ) en el que se halla definida una probabilidad P. Ω representa "lo que puede ocurrir en el mundo" (aparece una nueva reserva de petróleo, Bin Laden ataca las torres gemelas, Bill Gates decide respetar la ley anti-monopolio, etc). Parece claro que será muy difícil (más bien imposible) obtener información útil acerca de la probabilidad P. Afortunadamente para nosotros, veremos que lo único que usamos de P en los cálculos es quiénes son sus conjuntos de medida 0, es decir, admitiremos que sabemos distinguir los sucesos imposibles de los posibles, aunque no sepamos asignarles a estos últimos una probabilidad.

En general, supondremos por simplificar que Ω es finito, lo que hace que tengamos menos problemas "técnicos" (medibilidad, finitud de las normas, etc). En ese caso, habitualmente asumimos que todos los elementos de Ω tienen probabilidad estrictamente mayor que 0, ya que eliminamos de Ω los sucesos imposibles. Una observación, útil en ocasiones: es equivalente pensar que Ω es finito a pensar que P es puramente atómica, con una cantidad finita de átomos.

Veamos cómo se incluye en el modelo la información. La idea es que uno es capaz de conocer las funciones *medibles*, y mediante estas tenemos que ser capaces de decidir, por ejemplo, si un determinado suceso ha ocurrido o no. Es decir, dado un determinado suceso A, queremos evaluar χ_A, o algo similar, que tome valores distintos en A y en A^c. Pero que χ_A sea medible (respecto de cierta σ-álgebra) es equivalente a pedir que A pertenezca a dicha σ-álgebra. Por tanto, resulta claro que la información va implícita en la σ-álgebra, y que σ-álgebras más finas portan más información que las σ-álgebras menos finas. En uno de los extremos tenemos la σ-álgebra trivial $\{\emptyset, \Omega\}$ y en el otro $\mathcal{P}(\Omega)$. Las funciones medibles sobre la σ-álgebra trivial son sólo las constantes, y por tanto ahí no es posible distinguir más que Ω de $\emptyset$. En cambio, con (las funciones medibles respecto de) $\mathcal{P}(\Omega)$ podemos distinguir dos puntos cualesquiera de Ω.

Esa situación es estática, pero a nosotros nos va a interesar incluir en nuestro modelo el "crecimiento de la información". En $t = 0$ se sabe algo, en $t = 1$ ha pasado el tiempo y se sabe más, en $t = 2$ se sabe más aún, etc.

A la vista de los comentarios anteriores, parece natural considerar una cadena de σ-álgebras

$$\Sigma_0 \subset \Sigma_1 \subset \Sigma_2 \subset \cdots \subset \Sigma_T,$$

es decir, una filtración. En cada instante n, Σ_n representa la información disponible a los agentes en ese momento.

Habitualmente, en $t = 0$, es decir "ahora", consideramos la situación *determinista*, es decir, conocemos los precios, las cotizaciones, etc. Por tanto no hay incertidumbre y eso hace que $\Sigma_0 = \{\emptyset, \Omega\}$. Al otro lado, si nuestro horizonte temporal llega hasta T, no nos va a interesar refinar más allá de lo que se sepa en T. Lo que hacemos por tanto es considerar Σ_T como $\mathcal{P}(\Omega)$, es decir, identificamos los puntos que en $t = T$ no sean distinguibles.

Para mí, la forma más natural de entender esto es pensar en una estructura de árbol, y qué mejor estructura arbórea que el desarrollo (decimal, binario, etc) de, por ejemplo, un número real. Si lo pensamos en binario, encajará perfectamente con el modelo binomial CRR (Cox-Ross-Rubinstein). Identifiquemos Ω con el conjunto de los numeros que admitan un desarrollo binario de, por ejemplo, 10 cifras. En cada nuevo instante vamos conociendo una nueva cifra del desarrollo. Así, tras $n = 1$ sabemos distinguir si nuestro número empieza por 0 o por 1, tras $n = 2$ podemos distinguir si empieza por 00, 01, 10 o 11, etc. Identificad como ejercicio en este modelo la filtración (Σ_n).

2. El proceso de precios

Dentro de esta estructura de información, tenemos que ubicar la información propiamente dicha. Nuestra información básica es el *proceso de precios* $S = (S_n)$. Supondremos que existen una cantidad d finita de activos cuya cotización varía y un activo no determinista, denominado *activo sin riesgo* (riskless), que en principio corresponde a un bono del tesoro, o dinero en el banco, o equivalentes. Esto nos da un total de $d+1$ activos cuyo valor en cada instante t en función de $\omega \in \Omega$ modelizamos por medio de un proceso estocástico S con valores en $\mathbb{R}^{d+1}$ (o, equivalentemente, con $d+1$ procesos estocásticos reales S^i con $0 \leq i \leq d$). La notación es

$$S = (S_n^i(\omega))_{n=1,i=0}^{T,\quad d},$$

donde $S_n^i(\omega)$ representa el valor del activo i en el instante $t = n$ en el "estado del mundo" ω. Típicamente el activo 0-ésimo es el activo sin riesgo.

Puesto que necesariamente los valores actuales sólo pueden depender de la información disponible *ahora*, no de la disponible mañana, suponemos que S es un proceso *adaptado*.

El activo sin riesgo juega el papel del *numerario* (numeraire), la unidad monetaria a la que todo está referida. Parece más o menos evidente que si de cualquier otro de los activos sabemos que su valor es siempre (en todo n y en todo ω) estrictamente positivo, podríamos considerar éste como activo sin riesgo, y referirlo todo a su valor; es decir, podríamos valorar las cosas por ejemplo en gramos de oro, en lugar de en euros, pero esto obviamente haría que el euro fuera un activo con riesgo.

El activo sin riesgo habitualmente aumenta de valor en el tiempo, debido al *tipo de interés*. Es decir, no depende de ω, pero sí de n. Veremos más adelante que las cuentas son mucho más sencillas y las ideas más claras si uno trabaja sin tipos de interés. Hay dos formas, equivalentes, de hacer esto. Una de ellas es trabajar suponiendo que el tipo de interés es 0, es decir que el dinero tiene valor constante. La otra es trabajar con *valores descontados*. La idea sería dar los precios en "euros de $t = 0$". Esto es muy sencillo, y consiste simplemente en dividir S_n^i $(1 \leq i \leq d)$ por S_n^0, de manera que los euros "siempre valen 1". Al proceso de *precios descontados* así obtenido lo llamaremos $\overline{S}$, y es con el que realmente trabajaremos.

3. El comportamiento de los agentes

Necesitamos ahora modelizar el comportamiento de los distintos agentes económicos. Un agente económico es alguien que, en cada instante temporal puede comprar y/o vender activos (o derivados) del mercado, y que en función del tiempo va modificando estas compras. Por lo tanto, lo que nos interesa de el en cada momento n es su *cartera* (portfolio)

$$\theta_n = (\theta_n^0, \theta_n^1, \dots, \theta_n^d),$$

donde θ_n^i representa la cantidad del activo i en posesión del agente en el instante n. Suponemos que θ^i puede tomar valores negativos, ya que uno puede vender activos que no se poseen.

Claramente el comportamiento de los agentes no es determinista: en función de los diferentes estados del mundo, tomarán unas decisiones u otras que resultarán en diferentes carteras. Por ello consideramos $\theta = (\theta_n)$ como un proceso estocástico, es decir, θ_n depende también de ω.

Asumiremos que la dinámica es la siguiente: una vez que los precios y el estado del mundo en $t = n - 1$ se conocen, el agente toma sus decisiones (compra, vende, etc.) lo que lleva a que configure la cartera θ_n que mantendrá hasta el instante n, en que los precios volverán a cambiar y ella configurará su cartera θ_{n+1}, etc. Es decir, asumimos que el proceso θ es *predecible*. Obsérvese en particular que θ_0 no está definido. Al proceso θ también le llamaremos *estrategia*, puesto que describe precisamente la estrategia económica que sigue el agente.

4. El proceso de valor y el proceso de ganancias

Una vez que tenemos definidos θ y S es natural definir dos procesos asociados a estos. Uno es el proceso de valor, V, definido como

$$V_n = \begin{cases} \theta_1 \cdot S_0 = \displaystyle\sum_{i=0}^{d} \theta_1^i S_0^i & \text{si } n = 0 \\ \\ \theta_n \cdot S_n = \sum_{i=0}^{d} \theta_n^i S_n^i & \text{si } n > 0 \end{cases}$$

V_n nos indica el valor (aleatorio) de la cartera θ en el instante n. El pequeño ajuste necesario en la definición se debe que en $n = 0$ no hay cartera.

Está implícita en la definición del proceso V una de nuestras últimas hipótesis sobre el mercado: suponemos que el mercado no tiene fricciones (frictionless), lo que quiere decir que el precio de compra y de venta de los activos es el mismo. Esto hace que en el cálculo de V no nos tengamos que preocupar de si los valores de θ_n^i son positivos o negativos, o, si se prefiere, la misma hipótesis va implícita en el hecho de que haya un único proceso de precios (y no dos, uno de compra y otro de venta)

El otro proceso que nos queda por definir es el proceso de ganancias, G. Con la notación de la Definición 5.1, G es la transformada de S por θ, de manera que definimos

$$G_n = \begin{cases} 0 & \text{si } n = 0 \\ \sum_{m=1}^{n} \theta_m (S_m - S_{m-1}) = \sum_{m=1}^{n} \sum_{i=0}^{d} \theta_m^i (S_m^i - S_{m-1}^i) \\ \quad \text{si } n > 0 \end{cases}$$

5. Aproximándose a la noción de arbitraje: las estrategias

Buscamos introducir en nuestro modelo la noción de arbitraje en forma matemática, de manera que podamos obtener consecuencias de pedir que el modelo esté libre de arbitraje.

Como queda implícito en el primer capítulo, un arbitraje es una *estrategia* que nos permita ganar dinero sin arriesgar. Por lo tanto, para poder introducir la noción de arbitraje en nuestros razonamientos, necesitamos primeramente definir con claridad cuáles son las estrategias que pueden seguir los agentes.

Una estrategia θ se dice *autofinanciada*, si no se introduce ni se extrae dinero adicional después de $t = 0$. Es decir, si los cambios en el valor V se deben únicamente a las variaciones en las cotizaciones de los activos. Teniendo en cuenta que los activos cambian de valor en los instantes n, antes de modificarse las carteras, lo que esto quiere decir es que el agente, a la vista de lo ocurrido en n, decide modificar su cartera (vender ciertos activos y comprar otros) sin introducir ni retirar dinero adicional.

Esto se traduce en que, para todo $1 \leq n \leq T - 1$,

$$\theta_{n+1} \cdot S_n = \theta_n \cdot S_n.$$

Utilizando la notación $\Delta X_n = X_n - X_{n-1}$, esto implica que

$$\Delta V_n = \theta_n \cdot S_n - \theta_{n-1} \cdot S_{n-1} = \theta_n \cdot S_n - \theta_n \cdot S_{n-1} = \theta_n \cdot \Delta S_n.$$

Se sigue inmediatamente de la definición del proceso de ganancias G que θ es autofinanciada si y sólo si

$$V_n = V_0 + G_n,$$

es decir, el valor de la cartera en un instante n se debe únicamente al valor de la inversión inicial más el rendimiento (positivo o negativo) de la cartera.

Veremos que en realidad nos va interesar habitualmente pensar en "euros constantes", es decir, suponer corregidos los valores de los activos de manera que el activo sin riesgo tenga valor constante. La manera más sencilla de hacer esto es considerar los procesos *descontados* (discounted) respectivos:

Definimos $\beta_n = \frac{1}{S_m^0}$, y consideramos el proceso de precios descontados como

$$\overline{S_n} = \beta_n S_n.$$

Análogamente, definimos el valor descontado y la ganancia descontada de una cartera como

$$\overline{V_n} = \beta_n V_n \quad \text{y}$$

$$\overline{G_n} = \beta_n G_n$$

6. Arbitraje

Ya tenemos a nuestro alcance todos los elementos necesarios para formalizar matemáticamente la noción de arbitraje, de manera que podamos trabajar con ella.

Definición 6.1. Un *arbitraje* es una estrategia autofinanciada θ que verifica

$$\begin{cases} V_0(\theta) = 0 \\ V_T(\theta) \geq 0, \text{ es decir, para (casi) todo } \omega \in \Omega \quad V_T(\theta)(\omega) \geq 0 \\ \mathbf{E}(V_T(\theta)) > 0 \end{cases}$$

Es decir, un arbitraje es una estrategia que con una inversión 0 en $t = 0$ garantiza que seguro que no se pierde dinero y que con probabilidad estrictamente mayor que 0 se ganará dinero. Es decir, es la posibilidad de ganar dinero sin arriesgar.

7. Mercados completos y valoración por arbitraje

Resulta inmediato cómo utilizar la noción de arbitraje para valorar derivados financieros. Para nosotros, un derivado será una función $H \in L_0(\Omega, \Sigma, P)$, es decir, una función medible que toma distintos valores (lo que paga el derivado, su "pay-off") en función de los posibles estados del mundo $\omega \in \Omega$. En la literatura se suele denominar a H un *contingent claim*. La idea para valorar derivados utilizando la noción de arbitraje es esencialmente la misma que usamos en el sencillo ejemplo del Capítulo 2: buscamos *replicar* ("hedge") H por medio de una estrategia $\theta = (\theta_n)$ que cuyo valor asociado $V = (V_n)$ coincida en T con H, es decir, tal que $V_T(\omega) = H(\omega)$ para (casi) todo $\omega \in \Omega$.

Supongamos que H es *replicable*, es decir, que existe un θ como en el párrafo anterior. Resulta claro que si $V_T = H$ *y suponemos la hipótesis de ausencia de arbitraje* entonces el valor de H *ahora* (es decir, en $t = 0$) debe de ser V_0 (que por ser $\Sigma_0 = \{\emptyset, \Omega\}$-medible es constante). Cualquier otra posibilidad nos llevaría a un arbitraje, como en el ejemplo del Capítulo 2.

Análogamente, es fácil demostrar que si existen dos estrategias admisibles θ y ϑ tales que sus respectivos valores en T, $V_T(\theta)$ y $V_T(\vartheta)$, coinciden, entonces $V_0(\theta) = V_0(\vartheta)$, de manera que el precio de H queda unívocamente determinado.

EJERCICIO 7.1. *Probar las dos últimas afirmaciones.* **Sugerencia:** *En ambos casos, vamos a suponer que el resultado es falso, y a partir de ahí construimos un arbitraje. En el primer caso, suponemos que el precio de H fuera distinto de V_0, consideramos H como un activo más y construimos la estrategia $(\theta, -1)$ (o $(-\theta, 1)$, formada por θ y una unidad de H.*

Para la segunda afirmación, suponemos que $V_0(\theta) \neq V_0(\vartheta)$, y consideramos la cartera $\theta - \vartheta$.

Una pregunta natural ahora es si todo contingent claim es replicable. Los mercados en los que esto ocurre se llaman *completos.* Habitualmente los mercados son incompletos, y aún en el caso de que un derivado fuera replicable, no resulta nada obvio como construirle una estrategia replicante (hedging strategy). El Teorema Fundamental de Valoración de Activos nos permitirá valorar derivados sin necesidad de conocer una estrategia replicante.

Capítulo 4

El Teorema Fundamental de Valoración de Activos

Demostraremos en este capítulo el resultado principal de este minicurso, en el caso de tiempo discreto y finito y Ω finito.

1. El Teorema

Suponemos nuestro mercado definido mediante una base estocástica $(\Omega, \Sigma, P, \Sigma_n, \{0, \ldots, T\})$, y un proceso de precios $S = (S_n)$ definidos como en el capítulo anterior. En particular, recordemos que Ω es finito, que todos sus elementos tienen probabilidad estrictamente positiva, que S toma valores en $\mathbb{R}^{d+1}$ y que su primera coordenada $S^0 = (S^0)_n$ representa el activo sin riesgo. Recordemos también que podemos definir el proceso de precios descontados $\overline{S} = (\overline{S})_n$ sin más que dividir S_n por S_n^0.

Necesitaremos una definición auxiliar.

Definición 1.1. En las condiciones anteriores, una probabilidad Q sobre (Ω, Σ) se dice una *medida martingala equivalente* (equivalent martingale measure, EMM) si Q es equivalente a P (es decir, ambas tienen los mismos conjuntos de medida nula) y además $\overline{S}$ es una martingala respecto de Q, es decir, para cada $0 \leq i \leq d$, $\overline{S}^i$ es una Q-martingala.

Ahora ya podemos enunciar el Teorema Fundamental de Valoración de Activos (para el caso de tiempo discreto y finito y Ω finito).

Teorema 1.2. *Sea un mercado definido como en los párrafos anteriores. Este mercado está libre de arbitraje si y sólo si existe una medida martingala equivalente Q.*

Observación 1.3. *Antes de hacer la demostración del teorema, haremos un comentario que simplifica la demostración. Llamemos $\theta', \overline{S}'$ y $\overline{G}'(\theta')$ a los procesos resultantes de eliminar la primera coordenada ($\theta^0, \overline{S}^0, \overline{G}^0$) de los respectivos procesos. Puesto que, por definición, el proceso $\overline{S}^0$ es constante, se sigue de las definiciones que*

$$\overline{G}'(\theta') = \overline{G}(\theta).$$

Además, necesitamos también observar que, dado un proceso predecible $\theta' = (\theta^1, \ldots, \theta^d)$ *cualquiera, existe un único proceso* θ^0 *que hace que la estrategia* $\theta = (\theta^0, \theta^1, \ldots, \theta^d)$ *sea una estrategia predecible,* autofinanciada *y que verifica* $V_T(\theta) = 0$. *La idea es simplemente utilizar* θ^0 *para compensar los cambios en el valor de* θ' *que no se deban a los cambios de los precios. Basta con definir* θ_1^0 *como el único valor que verifica*

$$\theta_1^0 S_0^0 = -\sum_{i=1}^d \theta_1^i S_0^i$$

y por inducción θ_{n+1}^0 *es la única variable* Σ_n*-medible que verifica*

$$\theta_{n+1}^0 S_n^0 = \theta_n^0 S_n^0 + \sum_{i=1}^d S_n^i(\theta_n^i - \theta_{n+1}^i)$$

.

Demostración. (Del Teorema Fundamental de Valoración de Activos) Una de las implicaciones es muy sencilla. Supongamos que existe una EMM Q. Es decir, $\overline{S}$ es una martingala respecto de una cierta probabilidad Q equivalente a P. Esto quiere decir que, para todo $n \geq 1$,

$$\mathbf{E}_Q(\Delta \overline{S}_n | \Sigma_{n-1}) = 0.$$

Sea θ una estrategia autofinanciada con valor inicial 0. Puesto que θ es autofinaciada, $\Delta V_n = \theta_n \Delta S_n$. De ahí se sigue inmediatamente que

$$\Delta \overline{V}_n = \theta_n \Delta \overline{S}_n = \sum_{i=0}^d \theta_n^i \Delta \overline{S}_n^i$$

Puesto que $\overline{S}$ es una Q-martingala, y θ es predecible, el Teorema 5.4 nos garantiza que $\Delta \overline{V}_n$ es una suma finita de Q-martingalas, y por tanto una Q-martingala él mismo.

Por tanto, usando el hecho de que las funciones Σ_0-medibles son constantes, se tiene

$$\mathbf{E}_Q(\overline{V}_T(\theta)) = \mathbf{E}_Q(\overline{V}_0(\theta)) = \mathbf{E}_P(\overline{V}_0(\theta)) = V_0(\theta) = 0.$$

Veamos ahora que no puede haber arbitrajes. Sea θ una estrategia autofinanciada con valor inicial 0, y supongamos además que $V_T(\theta) \geq 0$ en P-c.t.p. En ese caso, también $V_T(\theta) \geq 0$ en Q-c.t.p. Como además tenemos que $\mathbf{E}_Q(\overline{V}_T(\theta)) = 0$, se sigue que $V_T(\theta) = 0$ en Q-c.t.p., y por tanto también $V_T(\theta) = 0$ en P-c.t.p., luego no hay arbitrajes.

La otra implicación es más interesante. Supongamos que el mercado está libre de arbitraje. Puesto que Ω es finito y tiene, por ejemplo, m elementos, identificamos de manera natural $L_\infty(\Omega)$ (ó $L_p(\Omega)$ con el p

que prefiráis) con $\mathbb{R}^m$ (con la norma p). Sea $C \subset \mathbb{R}^m$ el cono positivo del que retiramos el vector 0, es decir,

$$C = \{(x_1, \ldots, x_m);\ \text{con } x_i \geq 0\ (1 \leq i \leq m) \text{ y } x_j > 0 \text{ para algún } j\}.$$

Sea θ una estrategia predecible y autofinanciada, y sea $\overline{G}(\theta)$ su proceso de ganancias descontadas.

<u>*Aserto*</u>: En la hipótesis de ausencia de arbitraje (NA), $\overline{G}_T(\theta) \notin C$

Demostración del Aserto: Supongamos inicialmente que $V_0(\theta) = 0$. En ese caso, (NA) implica que $V_T(\theta) \notin C$. Puesto que θ es autofinanciada, se sigue que $G_T(\theta) + V_0(\theta) = V_T(\theta)$, de donde se sigue lo pedido.

Supongamos ahora una estrategia cualquiera autofinanciada θ con valor inicial $V_0(\theta) \neq 0$. Consideramos la estrategia $\tilde{\theta}$ asociada definida para cada n como

$$\tilde{\theta} = \begin{cases} \theta_n^0 - V_0(\theta) \text{ si } i = 0 \\ \theta_n^i \text{ si } i \neq 0 \end{cases}$$

Es claro que $\tilde{\theta}$ es autofinanciada y que $V_0(\tilde{\theta}) = 0$. Además, mirando la definición de $\overline{G}$, se sigue inmediatamente que, para todo n,

$$\overline{G}_n(\theta) = \overline{G}_n(\tilde{\theta}),$$

ya que $\Delta \overline{S}_n^0 = 0$.

Por lo tanto, puesto que, por el caso inicial, $G_T(\tilde{\theta}) \notin C$, se concluye la demostración del Aserto.

Veamos ahora cómo concluir la demostración del teorema.

Sea

$$L = \{\overline{G}_T(\theta);\ \text{donde } \theta \text{ es una estrategia predecible}\}.$$

Por la observación 1.3, se tiene que

$$L = \{\overline{G'}_T(\theta');\ \text{donde } \theta' \text{ es una estrategia predecible}\}.$$

Claramente L es un subespacio vectorial, y del Aserto se sigue que $L \cap C = \emptyset$. Querríamos aplicar algún Teorema de separación, pero todavía no podemos ya que C no es compacto, ni siquiera cerrado. Consideramos entonces

$$K = \{x \in C \text{ tales que } \mathbf{E}_P(x) = 1\}.$$

Ahora si, K es compacto, convexo, y $K \cap L = \emptyset$. Por tanto, los teoremas de separación tipo Hahn-Banach (puesto que estamos en el caso finito dimensional basta con alguna de las formas del Teorema de

Minkowski) nos garantizan que existen $x^* \in (\mathbb{R}^m)^*$, $a \in \mathbb{R}$, tales que, para todos $x \in K$, $y \in L$

$$0 = x^*(y) < a < x^*(x)$$

Escribimos x^* como

$$x^* = (p_1^*, \ldots, p_m^*).$$

Para todo $1 \leq i \leq m$, sea $\xi_i = (0, \ldots, 0, \frac{1}{p_i}, 0, \ldots, 0)$, con $P(\omega_i) = p_i$. Entonces $\mathbf{E}(\xi_i) = 1$, es decir, $\xi_i \in K$ y por tanto $x^*(\xi_i) = \frac{p_i^*}{p_i} > 0$. Por lo tanto $p_i^* > 0$.

Sea ahora

$$\alpha = \sum_{i=1}^{m} p_i^* \text{ y sea } y^* = \frac{x^*}{\alpha}.$$

Entonces

$$y^* = (q_1, \ldots, q_m),$$

con $q_i = \frac{p_i^*}{\alpha}$. Por lo tanto se tiene que

$$\sum_{i=1}^{m} q_i = 1$$

y podemos definir una probabilidad Q sobre Ω como

$$Q(\omega_i) = q_i.$$

Por construcción se tiene que Q es equivalente a P.

Para todo $x \in L$, se tiene que $y^*(x) = \frac{x^*(x)}{\alpha} = 0$. Además, claramente $y^*(x) = \mathbf{E}_Q(x)$. Por lo tanto, si θ' (con la notación de la observación 1.3) es una estrategia predecible, se tiene que $\overline{G}_T(\theta') \in L$ y entonces $\mathbf{E}_Q(\overline{G}_T(\theta') = 0$ para cualquier proceso predecible θ'. Ahora basta aplicar la Proposición 5.6 para obtener que $\overline{S'}$ es una martingala. Como $\overline{S}^0$ es constante, también es trivialmente una martingala, y ya podemos concluir que $\overline{S}$ es una martingala. $\square$

OBSERVACIÓN 1.4. *De las definiciones se sigue que ser una martingala, o ser una medida martingala, es algo que, en el caso discreto, se verifica "paso a paso", por ejemplo un proceso es una martingala si y sólo si cumple cierta condición en cada uno de los "saltos" de n a $n+1$. Está implícito en la demostración del Teorema, en realidad implícito de la demostración de la Proposición 5.6, que, en el caso discreto, también la existencia o no de arbitraje es una cuestión que se decide "paso a paso". En concreto, uno puede definir de forma natural lo que sería un* arbitraje en el paso n, *y demostrar que un mercado está libre de arbitraje si y sólo si está libre de arbitraje en cada uno de* *los pasos de n. Eso nos lleva a una construcción "paso a paso" de la medida martingala. Podéis leer acerca de esto en los libros* [**4, 7**]. *Lo que quisiera que en cualquier caso os quedara ahora más o menos claro es que, para el modelo discreto con horizonte finito, el caso de sólo dos instantes temporales* $0, 1$, *es, a muchos efectos, el caso más general.*

2. Valoración por arbitraje

Veamos cómo se puede utilizar el TFVA para valorar un activo, lo que de paso nos llevará a interpretar la medida martingala Q que aparece en el Teorema.

Supongamos un derivado H. Como en la Sección 7, supongamos que H fuera replicable por medio de una estrategia autofinanciada θ. Es decir, $V_T(\theta) = H$ o, lo que es lo mismo

$$\overline{V}_T(\theta) = \frac{H}{S_T^0}.$$

Puesto que $\overline{S}$ es una Q-martingala, también lo es $\overline{V}(\theta) = \overline{G}(\theta) + V_0(\theta)$. De ahí se sigue que, para todo n,

$$\overline{V}_n(\theta) = \mathbf{E}_Q(\overline{V}_T(\theta)|\Sigma_n) = \mathbf{E}_Q(\frac{H}{S_T^0}|\Sigma_n).$$

En particular, se tiene que, si $P(H)$ es el único precio posible de H con el que no hay arbitraje, entonces

$$P(H) = V_0(\theta) = \mathbf{E}_Q(\frac{H}{S_T^0}|\Sigma_0) = \mathbf{E}_Q(\frac{H}{S_T^0}).$$

Es decir, calcular el precio de un derivado H no es más que integrar el valor descontado del derivado con respecto de la medida martingala Q.

Esto nos lleva a interpretar Q como una medida *neutral al riesgo* (*risk-neutral*) en el sentido siguiente: la probabilidad P del mundo es desconocida (e incognoscible). Cada agente percibe el mundo con una probabilidad subjetiva distinta, que tiene que ver con su aversión al riesgo, y en base a esa probabilidad y a su aversión al riesgo toma decisiones económicas. La medida Q se dice libre de riesgo por lo siguiente: supongamos un agente al que el riesgo le resulta indiferente, y supongamos que la medida de este agente es Q. Puesto que el riesgo le resulta indiferente, no distingue entre un pago constante de valor K y un derivado que verifique que $\mathbf{E}_Q(H) = K$, y ambas cosas para ella deben valer lo mismo. Es decir, si uno es neutral al riesgo, para calcular precios lo que uno hace es calcular esperanzas. Y eso es lo que hace Q.

Para este agente, el precio de un derivado debe de ser exactamente la esperanza matemática de su valor, ya , es decir, dos carteras con la misma esperanza matemática le resultan indiferentes.

Capítulo 5

Valoración de derivados

1. Árboles con varios pasos

La idea del árbol binomial se puede extender al caso en que tenemos un árbol binomial con varios pasos. Veámoslo en un árbol de dos pasos. Supongamos que el valor inicial del activo es S_0 y que en cada paso temporal de amplitud Δt puede o bien subir, en cuyo caso su valor se multiplica por $u > e^{r\Delta t}$ o bien bajar en cuyo caso su valor se multiplica por $d < e^{r\Delta t}$. Suponemos que los valores de u y d no cambian en los sucesivos nodos.

Usando los razonamientos anteriores podemos probar que la medida neutral al riesgo p es

$$p = \frac{e^{r\Delta t} - d}{u - d}$$

Entonces, para calcular el valor de la opción en el nodo inicial comenzamos calculando su valor en cada uno de los tres posibles nodos posibles. A continuación, calculando esperanzas con respecto de p calculamos los valores en cada uno de los nodos intermedios y finalmente, calculando de nuevo esperanzas con respecto de p (nótese que p es la misma para los dos saltos, puesto que u y d no varían), calculamos el valor de la opción en $t = 0$.

2. Martingalas en al árbol binomial

Obviamente, debería haber una relación entre los razonamientos elementales que hicimos en la construcción del árbol binomial y el Teorema Fundamental de Valoración de Activos. En efecto, es fácil ver que, bajo la probabilidad definida en el modelo elemental como probabilidad de subir igual a

$$p = \frac{(1 + r)S - S_d}{S_u - S_d}$$

y probabilidad de bajar igual a $1 - p$, el proceso de precios descontados se convierte en una martingala. Pensemos en el caso del árbol binomial de varios pasos definido antes, en el que subir supone multiplicar por un factor $u > 1 + r$ y bajar multiplicar por $d < 1 + r$. Recordemos que en ese árbol, si en un cierto nodo ω del nivel n el precio del subyacente es $S_n(\omega)$ entonces en los nodos descendientes inmediatos suyos, los precios son $S_n u$ y $S_n d$. Además, para cada n, Σ_n es la σ-álgebra mínima capaz de distinguir entre sí los nodos del nivel n. No es difícil ver (no escribimos las cuentas aquí, pero son simples) que en este modelo

$$\mathbf{E}(S_{n+1}|\Sigma_n)(\omega) = pS_n(\omega)u + (1-p)S_n(\omega)d.$$

También es fácil ver que, en este caso, la probabilidad de subir en cada nodo es siempre igual y vale

$$p = \frac{1+r-d}{u-d}$$

Por tanto, se tiene

$$\mathbf{E}(S_{n+1}|\Sigma_n)(\omega) = pS_n(\omega)u + (1-p)S_n(\omega)d =$$
$$= \frac{((1+r)-d)S_n(\omega)u}{u-d} + \frac{(u-(1+r))S_n(\omega)d}{u-d} = (1+r)S_n(\omega),$$

lo que implica que el proceso de precios descontados $\frac{S_n}{(1+r)^n}$ es una p-martingala.

3. Paridad Put-Call

Una *opción de venta* (*put-option*) es un instrumento que otorga el derecho, y no la obligación, de vender una unidad del subyacente a un precio de ejercicio prefijado en la fecha de vencimiento. Esa sería una opción de venta europea. Una opción de compra americana es igual salvo que la venta se puede hacer en cualquier momento de la vida de la opción. El valor a vencimiento de la put es $(K-S_T)^+ = \text{máx}(K-S_T, 0)$

Veamos que, para opciones europeas, los precios de una call y de una put están relacionados, de manera que, conocido uno de ellos, se puede conocer el otro.

Considérense las siguientes dos carteras: Por un lado, una opción call más Ke^{-rT} unidades monetarias en efectivo. Por otro lado, una opción put más una unidad del subyacente.

Es fácil ver que, a vencimiento, ambas carteras valen lo mismo, $\text{máx}(S_T, K)$. Puesto que a vencimiento ambas carteras valen lo mismo, y puesto que no es posible ejercer las opciones anticipadamente, ambas carteras deben valer lo mismo a día de hoy. Es decir,

$$C + Ke^{-rT} = P + S_0, \tag{1}$$

donde P y C son los precios de la Put y la Call respectivamente.

Ejercicio 3.1. *Comprobar la fórmula (1) mediante razonamientos de arbitraje.*

Por este motivo, no estudiaremos en ningún momento la valoración de opciones put europeas.

Para opciones americanas no existe una relación similar a la anterior, por lo que es necesario valorar separadamente las opciones put y call.

4. Opciones americanas sobre acciones que no pagan dividendos

Veamos que nunca es óptimo ejercer anticipadamente una opción de venta americana sobre una acción que no paga dividendos, de manera que una opción americana sobre este tipo de acción debe valer exactamente lo mismo que la opción europea.

Sea C_A y C_E los precios de una opción de venta americana y una europea respectivamente, ambas sobre el mismo activo (cuyo valor denotaremos por S), con el mismo strike K y el mismo tiempo de vencimiento T. Entonces se tiene

$$C_A \geq C_E \geq S_0 - e^{-rT}K \geq S_0 - K,$$

donde para la seguna desigualdad hemos utilizado la paridad Put-Call (junto con el hecho de que una Put nunca tiene valor negativo) y para la tercera hemos utilizado que los tipos de interés nunca son negativos.

De ahí se sigue que en el instante 0 no es óptimo de ejercer, ya que el precio de la opción americana (es decir, el precio al que podríamos venederla en el mercado) es superior al beneficio obtenido en caso de ejercer. Por supuesto, el razonamiento hecho en el $t = 0$ se prodría hacer en cualquier otro instante t, de donde se sigue que nunca será óptimo ejercer, por lo que la opción europea y la americana valen lo mismo.

Esta relación no se mantendrá en el caso de opciones sobre activos que paguen dividendos, como veremos más adelante.

5. Delta

Cuando veamos la versión en tiempo continuo de estos modelos, llamaremos *delta* (Δ) a la derivada del precio de la opción con respecto al precio del subyacente. Δ resulta muy importante a la hora de construir *carteras de cobertura*, noción que veremos más adelante.

Veamos ahora estas ideas en el caso de nuestro sencillo modelo. En el caso de un solo árbol binomial de un solo paso, en $t = 0$ Δ vale lo que allí llamábamos β,

$$\Delta = \frac{C_u - C_d}{S_u - S_d}.$$

Obsérvese que, así definida, Δ es la versión discreta de la derivada del precio de la opción con respecto al precio del subyacente, y que a la vez esa era precisamente la cantidad de unidades del subyacente que necesitábamos para construir una *cartera de cobertura* de la opción que nos anulara los riesgos de ésta. No es casual la coincidencia de ambas cantidades, como veremos más adelante en tiempo continuo.

La *cobertura delta* es fundamental, y la practican todos los bancos de inversión para protegerse del riesgo implícito en las opciones.

6. Movimiento Browniano: Introducción

Necesitamos definir movimiento Browniano, y establecer algunas de sus propiedades elementales, puesto que es parte esencial del modelo que nos vamos a creer para la cotización de una acción en tiempo continuo.

El movimiento Browniano recibe su nombre en honor a Robert Brown, botánico escocés del siglo XIX que, al estudiar granos de polen en suspensión observó que describían un movimiento aleatorio que le llevo a pensar que estaban vivos.

Necesitaremos la noción de movimiento Browniano, como un proceso estocástico que cumple ciertas propiedades, pero no es en absoluto trivial probar la existencia (y unicidad esencial) de tal proceso.

Definición 6.1. Un procesos estocástico $W = (W_t)_{t \geq 0}$ definido sobre (Ω, Σ, P) se dice que es un *movimiento Browniano* o un *Proceso de Wiener* si tiene trayectorias continuas e incrementos independientes estacionarios Gaussianos. Es decir, si

1. $W_0 = 0$ c.t.p.
2. La aplicación $t \mapsto W_t(\omega)$ es continua para casi todo $\omega \in \Omega$
3. Para todo $s \leq t$, $B_t - B_s$ es una variable aleatoria Gaussiana con media 0 y varianza $t - s$, y además $B_t - B_s$ es independiente de $\mathcal{F}_s := \sigma\{B_u; u \leq s\}$.

Como ya hemos dicho, el siguiente resultado no es en absoluto trivial. Su demostración excede el interés de este curso.

TEOREMA 6.2. *Existe un espacio de medida sobre el que se puede definir un movimiento Browniano.*

Dado que $N(0,\sigma) = \sigma N(0,1)$ y dado que $W_t - W_s$ está distribuido como una normal $N(0, \sqrt{t-s}) = \sqrt{t-s}N(0,1)$ escribiremos en notación diferencial

$$dW_t = \sqrt{dt}\, Z,$$

donde Z es una variable aleatoria distribuida como $N(0,1)$

Es conveniente para entender el movimiento Browniano que los *paseos aleatorios* (*random walks*) con saltos cada vez menores convergen al movimiento Browniano. Esta idea resulta muy útil a la hora de operar con el movimiento Browniano, ya que lo que se hace es aproximarlo discretamente por paseos aleatorios.

Veamoslo: Para todo $N \in \mathbb{N}$ suponemos N variables aleatorias independientes idénticamente distribuidas que pueden tomar los valores 1, con probabilidad $\frac{1}{2}$, y -1 con probabilidad $\frac{1}{2}$. Llamamos a estas variables $(Z_j)_{j=1}^N$. Es fácil ver que su esperanza y varianza son, respectivamente, 0 y 1

Nos va a interesar considerar la variable $X_N = \sum_{j=1}^N Z_j$

Cuando dejamos que N tienda a infinito, se tiene que la variable

$$\frac{\sum_{j=1}^N Z_j}{\sqrt{N}}$$

converge en distribución a una $N(0,1)$, por el Teorema Central del Límite.

Recordemos

TEOREMA 6.3 (Teorema Central del Límite). *Sea $(Z_i)_{i\in\mathbb{N}}$ una sucesión de variables aleatorias independientes idénticamente distribuidas con media μ y varianza σ^2. Entonces la variable aleatoria*

$$Y_N := \frac{\sum_{i=1}^N (X_i - \mu)}{\sigma\sqrt{N}}$$

converge en distribución, cuando N tiende a infinito, a una variable $N(0,1)$.

Dicho de otra manera, suponiendo momentáneamente $p = \frac{1}{2}$, la distribución de la suma $X_N = \sum_{i=1}^N Z_i$ converge en distribución a una normal de media 0 y varianza N

Queremos ver ahora cómo tenemos que adaptar las escalas para subdividir un intervalo temporal fijo T en una cantidad cada vez mayor

N de subintervalos de tamaño $\Delta t = \frac{T}{N}$. Si el "salto" de las Z_i no es 1 sino h, entonces su media y su varianza pasan a ser h y h^2 respectivamente.

En ese caso, aplicando el TCL se tiene que la variable

$$X_N = \sum_{i=1}^{N} Z_i$$

converge a una normal de media 0 y varianza $h^2 N$.

Vemos entonces que para que la varianza no sea 0 (sería un proceso determinista) ni infinita (tales procesos no modelizarían un proceso de precios) es necesario que h^2 sea del orden de $\frac{1}{N}$. En concreto, para converger a una normal $N(0, \sqrt{T})$ pedimos que $h^2 = \Delta t$, y ya veremos que el paseo aleatorio así construido converge a un movimiento Browniano.

7. El modelo de precios de una acción en tiempo continuo

Describimos en esta sección el modelo que asumiremos cierto para la cotización de una acción en tiempo continuo. En la sección siguiente propondremos una posible discretización de este modelo mediante árboles binomiales.

La propuesta inicial de Bachelier sugería que la aleatoriedad de los precios se podía explicar mediante un movimiento Browniano. Esta idea fue refinada por Samuelson en los años 60, quien sugiere que los precios siguen un proceso conocido como geometric brownian motion, o log-normal. Lo describimos a continuación.

Supongamos inicialmente que el proceso de precios S_t no tiene parte aleatoria, y que su apreciación (o depreciación) es constante en el tiempo. En ese caso es claro que

$$S_t = S_0 e^{\mu t}$$

o, en notación diferencial,

$$\frac{dS_t}{S_t} = \mu dt.$$

Este es el modelo que se utiliza para un activo sin riesgo.

La aleatoriedad se incluirá haciendo

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t,$$

donde W_t es un movimiento browniano.

Se sigue del Lema de Ito que, en ese caso,

$$S_t = S_0 e^{(\mu - \frac{\sigma^2}{2})t - \sigma W_t}.$$

Si definimos al rendimiento “continuo” del activo en el periodo $[0, t]$ como

$$\ln \frac{S_t}{S_0} = (\mu - \frac{\sigma^2}{2})t - \sigma W_t$$

podemos hacer los siguientes razonamientos que nos permiten intuir algunos de los comportamientos básicos de este modelo.

Notemos además que σ, la *volatilidad*, es un parámetro fácilmente observable en los mercados, mientras que μ, la *tendencia, deriva* o *drift* no lo es.

Entonces el rendimiento verifica

- Está distribuido como una normal de media 0 y varianza proporcional a Δt (o desviación típica proporcional a $\sqrt{\Delta t}$).
- Para todo $t > t_0$, $\ln \frac{S_t}{S_{t_0}}$ es independiente de S_{t_0}, y en general de todo lo sucedido antes de t_0.

El hecho de que los rendimientos están distribuidos como una normal es seguramente la parte más discutible de todo el modelo, y de hecho la parte que más a menudo se cambia cuando se intenta mejorar el modelo. Tiene la ventaja de que la normal es una distribución muy tratable, por lo que muchos razonamientos se pueden formalizar con relativa facilidad.

El principal inconveniente observado es que predice menos fenómenos extremos de los que realmente suceden en el mercado. En ciertos modelos se intenta mejorar este inconveniente añadiendo una Bernouilli, o cambiando la normal por distribuciones con *colas pesadas*, como una t de Student.

8. La construcción del modelo en el caso discreto

Veamos ahora qué sucede con nuestros árboles binomiales “multiplicativos”.

Tenemos un activo con precio inicial S_0 y un tipo de interï£¡s (anual, por ejemplo) r, expresado como interï£¡s continuo. S_n denotarï£¡ el valor del activo tras n pasos de “tamaï£¡o" Δt, de manera que $S_n e^{-rn\Delta t}$ es el valor del activo en euros descontados (en euros de $t = 0$).

Consideramos el árbol en el que el precio descontado del activo en cada ocasión puede, o bien subir, multiplicando por u, o bien bajar, multiplicando por d, con probabilidad $\frac{1}{2}$ en ambos casos. Llamamos Z_i a cada una de estas variables aleatorias. Entonces nuestro proceso ahora viene dado por la variable aleatoria

$$S_N = e^{rN\Delta t} S_0 \prod_{i=1}^{N} Z_i.$$

Para compararlo con el proceso en tiempo continuo que presentamos en la Sección 7, consideramos el proceso

$$\ln \frac{S_N e^{-rN\Delta t}}{S_0} = \ln \left(\prod_{i=1}^{N} Z_i \right) = \sum_{i=1}^{N} \ln Z_i$$

y vemos que cada una de las $\ln Z_i$ es una variable de Bernuilli que toma los valores $\ln u$ y $\ln d$ con probabilidades $\frac{1}{2}$ en ambos casos.

Claramente, la esperanza de $\ln Z_i$ vale $\frac{1}{2}(\ln u + \ln d)$. Su varianza es

$$\sigma_{\Delta t} = \mathbf{E}((\ln Z_i)^2) - (\mathbf{E}(\ln Z_i))^2 = \frac{1}{4}(\ln u - \ln d)^2.$$

Entonces, el Teorema Central del Límite nos dice que, cuando N tiende a infinito, $\ln \frac{S_N e^{-rN\Delta t}}{S_0}$ converge a una normal de media $\frac{N}{2}(\ln u + \ln d)$ y varianza $\frac{N}{4}(\ln u - \ln d)^2$. Buscamos ajustar adecuadamente u y d. Introducimos la restricción adicional $ud = 1$ y llegamos a que la media es 0 y la varianza $N(\ln u)^2$.

Queremos entonces ajustar u para que la volatilidad en el periodo Δt coincida con la volatilidad para este periodo predicha por el modelo en tiempo continuo que vimos en la Sección 7. Al final de dicha sección mencionamos que dicha volatilidad es $\sigma\sqrt{\Delta t}$. Por tanto, tenemos

$$\ln u = \sigma\sqrt{\Delta t},$$

es decir,

$$u = e^{\sigma\sqrt{\Delta t}}.$$

9. Valoración de opciones americanas usando árboles binomiales

La única diferencia entre una opción americana y una opción europea es que la opción americana da más derechos, puesto que se puede ejercer antes de vencimiento. Ya hemos visto que podemos valorar una opción europea de compra “resolviendo hacia atrás” en el árbol binomial, aplicando en cada nodo los razonamientos elementales para el caso de un único periodo de tiempo. A la hora de valorar una opción americana entonces hacemos lo siguiente:

Los pagos de la opción en los últimos nodos (es decir, a vencimiento) coinciden con los de la opción europea. Nos vamos entonces a los penultimos nodos (es decir, hemos dado $N-1$ pasos). En cada uno de esos nodos podemos hacer dos cosas: guardar la opción hasta vencimiento, en cuyo caso se convierte en una opción cuyos precios ya conocemos, o ejercer. Por lo tanto, el valor en estos nodos será el máximo de estas dos cantidades. De esta manera conocemos el valor de la opción en los nodos del nivel $N-1$. Con esta información podemos solucionar el nivel $N-2$, de nuevo tomando el máximo entre guardarla (para valorar eso de nuevo usamos la esperanza con respecto a la medida martingala) o ejercer. Retrocediendo de esta manera hasta el nodo 0 valoramos la opción.

10. Primera versión de la fórmula de Black-Scholes

Veamos el modelo de Cox-Ross-Rubinstein que nos permite entender la fórmula de Black-Scholes aproximando por medio de un modelo discreto (y por lo tanto sin la necesidad del cálculo estocástico de Itô del que hablaremos más adelante).

Suponemos una opción de compra con strike K, vencimiento T sobre un activo cuya cotización denotaremos por S. Subdividimos el intervalo $[0,T]$ en N intervalos iguales de tamaño $\frac{T}{N}$, y denotamos por S_n la cotización del activo en el instante $n\frac{T}{N}$.

La representamos mediante el árbol binomial antes descrito. Recordemos que valorar la opción será calcular la esperanza de sus pagos con respecto de la medida martingala.

La medida martingala la podemos obtener de forma única y obtenemos que la probabilidad de "subir" (u) es

$$p = \frac{1-d}{u-d}$$

y la de bajar d es

$$1-p == \frac{u-1}{u-d}.$$

El valor de la opción a vencimiento es como siempre $\max\{S_T - K, 0\}$. Tenemos que determinar S_T en función de $\omega \in \Omega$, donde Ω es el conjunto de los estados del mundo, que en nuestro caso podemos identificar con el conjunto de posibles caminos seguidos por la cotización S. Es fácil darse cuenta que S_T depende únicamente del número de subidas u y bajadas d que hayan tenido lugar, y no del orden en que estas hayan ocurrido.

Por lo tanto S_T es una variable aleatoria binomial con parámetros N y p, es decir, llamando Q a la medida martingala con respecto de la cual hemos de integrar se tiene

$$Q(S_T = u^j d^{N-j} S_0) = \binom{N}{j} p^j (1-p)^{N-j},$$

de manera que el valor C de la opción vendrá dado por

$$C = \sum_{j=0}^{N} \binom{N}{j} p^j (1-p)^{N-j} (u^j d^{N-j} S - K)^+$$

Podemos entonces utilizar los u_N y d_N calculados más arriba, valorar la opción para cada valor de N y a continuación dejar que N tienda a infinito. No presentamos las cuentas, por ser largas y tediosas. No conllevan ninguna dificultad grande más allá del uso adecuado del Teorema Central del Límite y de propiedades bien conocidas de las distribuciones normal y binomial. Se pueden consultar los detalles en [**4**] o [**7**]. El resultado al que se llega es la bien conocida y muy usada fórmula de Black-Scholes

$$C = S_0 \Phi(d_+) - e^{-rT} K \Phi(d_-)$$

donde C es el precio de la opción, K el precio de ejercicio, r el tipo de interés compuesto, T el tiempo, Φ es la función de distribución acumulada de la normal standard y

$$d_\pm = \frac{\ln\left(\frac{S_0}{K}\right) + (r \pm \frac{\sigma^2}{2})T}{\sigma\sqrt{T}}.$$

Queremos hacer notar que el precio de la opción es *independiente* de μ, la tendencia del valor de la acción, y en cambio es dependiente de σ, la volatilidad. El modelo de Black, Scholes y Merton da explicación matemática a este hecho, que los mercados ya intuían (es decir, lo reflejaban en los precios de las opciones) sin una justificación adecuada. Los trabajos de Black, Scholes y Merton fueron premiados con un premio Nobel en 1997.

11. Aparición de martingalas: versión discreta

Recordemos que los paseos aleatorios convergen al movimiento Browniano. Por tanto, una de las versiones discretas del proceso $dS = S(\mu dt + \sigma dW)$ es

$$\Delta S = S(\mu \Delta t + \sigma\sqrt{\Delta t})Z$$

donde Z es una Bernouilli que toma los valores $\{-1, 1\}$, cada uno con probabilidad $\frac{1}{2}$. Lo que queremos hacer entonces es pasar a la probabilidad neutral al riesgo q y ver qué proceso sigue S bajo esta probabilidad. Recordemos que

$$q = \frac{e^{r\Delta t} S_0 - S_d}{S_u - S_d} = \frac{e^{r\Delta t} S_0 - S_0(1 + \mu\Delta t - \sigma\sqrt{\Delta t})}{2 S_0 \sigma \sqrt{\Delta t}} =$$

$$= \frac{1}{2} + \frac{(e^{r\Delta t} - 1) - \mu\Delta t}{2\sigma\sqrt{\Delta t}} = \frac{1}{2} + \frac{(r-\mu)\sqrt{\Delta t}}{2\sigma} + O(\Delta t)$$

de manera que q y $1-q$ tienden a $\frac{1}{2}$ cuando Δt tiende a 0.

Consideramos entonces el proceso de precios bajo la medida neutral al riesgo. Llamamos $\tilde{Z}$ a la Bernouilli de valores $\{1, -1\}$ con probabilidades q y $1-q$ respectivamente y tenemos

$$\Delta S = S(\mu\Delta t + \sigma\sqrt{\Delta t}\tilde{Z})$$

Observemos que la media de $\tilde{Z}$ vale

$$\mathbf{E}(\tilde{Z}) = q - (1-q) = 2q - 1 = 2(\frac{1}{2} + \frac{(r-\mu)\sqrt{\Delta t}}{2\sigma} + O(\Delta t)) =$$
$$\frac{(r-\mu)\sqrt{\Delta t}}{\sigma} + O(\Delta t)$$

de modo que

$$\mathbf{E}(\sqrt{\Delta t}\tilde{Z}) = \frac{(r-\mu)\Delta t}{\sigma} + O((\Delta t)^{\frac{3}{2}})$$

Entonces para la varianza de $\sqrt{\Delta t}\tilde{Z}$ tenemos

$$\mathbf{V}(\sqrt{\Delta t}\tilde{Z}) = \mathbf{E}((\sqrt{\Delta t}\tilde{Z} - \mathbf{E}(\tilde{Z}))^2) = \mathbf{E}\left(\left(\sqrt{\Delta t}\tilde{Z} - \frac{r-\mu}{\sigma}\Delta t\right)^2\right) =$$
$$= \Delta t(q + (1-q)) - \left(\frac{\mu - r}{\sigma}\right)^2 (\Delta t)^2 = \Delta t + O((\Delta t)^2).$$

Obsérvese que, al pasar al límite, para el cálculo de la varianza el valor de q no ha sido relevante (el valor de q sí se usa en el cálculo de $\mathbf{E}(\sqrt{\Delta t}\tilde{Z})$, pero este valor queda multiplicado por $(\Delta t)^2$ y desaparece en el paso al límite).

Por ello, llegamos a que la variable $\sqrt{\Delta t}\tilde{Z}$ tiene la misma media y varianza que la variable

$$\frac{r-\mu}{\sigma}\Delta t + \sqrt{\Delta t}Z$$

Recordemos que en el Teorema Central del Límite lo único relevante a la hora de estudiar la convergencia de una suma de variables aleatorias i.i.d es su media y su varianza. Por tanto sustituimos las variables Z por las $\tilde{Z}$ y llegamos a que el proceso discreto, con la medida neutral al riesgo

$$\Delta S = S(\mu\Delta t + \sigma\sqrt{\Delta t}\tilde{Z})$$

convergerá, cuando Δt tienda a 0, al mismo proceso al que converja

$$\Delta S = S\left(\mu\Delta t + \sigma\left(\frac{r-\mu}{\sigma}\Delta t + \sqrt{\Delta t}Z\right)\right) = S(r\Delta t + \sigma\sqrt{\Delta t}).$$

Por tanto, si S sigue el proceso

$$dS = S(\mu dt + \sigma dW)$$

bajo la probabilidad real, entonces el proceso correspondiente bajo la probabilidad neutral al riesgo es

$$dS = S(rdt + \sigma dW),$$

de manera que para pasar de la probabilidad real a la probabilidad neutral al riesgo basta con sustituir μ por r en el proceso de precios, sin modificar σ.

Ahora, el proceso de precios descontado es

$$dS = S\sigma dW,$$

y es inmediato comprobar que es una martingala.

12. El modelo

Como ya hemos dicho anteriormente, el modelo que aceptaremos para la evoluciï£¡n de los precios de un activo es, en notación diferencial,

$$dS = S(\mu dt + \sigma dW).$$

La demostración (e incluso el enunciado preciso) del Teorema Fundamental de Valoración de Activos en el caso de tiempo continuo escapa con mucho a los objetivos de este curso. Los lectores interesados la pueden leer en [**?**]. A nosotros nos bastará saber que el teorema sigue siendo esencialmente cierto, de manera que, esencialmente, se puede caracterizar la ausencia de arbitraje por medio de la existencia de una medida bajo la cual el proceso de precios descontado es una martingala. Sigue también siendo cierto que para valorar un derivado (al menos uno de tipo europeo) basta con integrar sus pagos a vencimiento respecto de la medida martingala.

13. Lognormal

Veamos la propiedad que justifica el nombre de lognormal para los procesos de precios de la forma $dS = S(\mu dt + \sigma dW)$.

Notemos primero que si aplicamos el Lema de Itô a una función $f(S,t)$. Se tiene

$$(2) \qquad df = \left(\frac{\partial f}{\partial S} \mu S + \frac{\partial f}{\partial t} + \frac{1}{2} \frac{\partial^2 f}{\partial S^2} \sigma^2 S^2 \right) dt + \frac{\partial f}{\partial S} \sigma S dW$$

Sea $f = \ln S$. Aplicando (2) tenemos que

$$df = (\mu - \frac{\sigma^2}{2})dt + \sigma dW.$$

Es decir, el logaritmo del proceso de precios sigue un proceso de Wiener generalizado de drift $\mu - \frac{\sigma^2}{2}$ y varianza σ^2. En particular, el cambio en la función $\ln S$ entre $t = 0$ y $t = T$ sigue una distribución normal de media $(\mu - \frac{\sigma^2}{2})T$ y varianza $\sigma^2 T$, es decir $\ln S_t - \ln S_0 = \ln \frac{S_T}{S_0}$ está distribuido como

$$(3) \qquad N((\mu - \frac{\sigma^2}{2})T, \sigma\sqrt{T}),$$

y por tanto $\ln S_t$ está distribuido como

$$N(\ln S_0 + (\mu - \frac{\sigma^2}{2})T, \sigma\sqrt{T}).$$

Es decir, $\ln S_T$ sigue una distribución normal.

A partir de aquí se sigue que

$$\ln \frac{S_t}{S_0} \sim (\mu - \frac{\sigma^2}{2})t + N(0, \sigma\sqrt{t})$$

y por tanto

$$S_t = S_0 e^{(\mu - \frac{\sigma^2}{2})t + \sigma\sqrt{t}},$$

y esta es la expresión, en notación no diferencial, del proceso de precios.

14. El rendimiento esperado

Al ver un proceso de precios de la forma $dS = S(\mu dt + \sigma dW)$, el primer impulso nos lleva a creer que el rendimiento esperado de dicho proceso será μ. Veamos que esto no es así.

Calculemos directamente su rendimiento medio. Ya hemos visto que el proceso se puede escribir como

$$S_t = S_0 e^{(\mu - \frac{\sigma^2}{2})t + \sigma W_t}.$$

Entonces, su rendimiento medio en un período Δt es

$$\mathbf{E}(\ln \frac{S_{\Delta t}}{S_0}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} ((\mu - \frac{\sigma^2}{2}\Delta t) + \sigma x) e^{-\frac{x^2}{2}} dx = (\mu - \frac{\sigma^2}{2})\Delta t$$

Otra forma de verlo es la siguiente. El rendimiento (expresado en forma continua) del activo es un número x que verifique

$$S_T = S_0 e^{xT},$$

de manera que

$$x = \frac{1}{T} \ln \frac{S_T}{S_0}.$$

De (3) se sigue que x está distribuido como

$$N(\mu - \frac{\sigma^2}{2}, \frac{\sigma}{\sqrt{T}}),$$

de manera que efectivamente vemos de nuevo que el rendimiento medio no es μ sino $\mu - \frac{\sigma^2}{2}$.

15. Proceso de precios de los futuros

Ya se ha visto como valorar un futuro. Queremos ahora mostrar cómo podemos usar el Lema de Itô para modelizar el proceso de precios de un futuro. Aparte del interés que esto tiene en sí mismo, es necesario para posteriormente valorar derivados sobre futuros.

Consideramos un contrato de futuros sobre un activo que no paga dividendos cuyo precio actual es S_0 y cuyo proceso de precios es $dS = S(\mu dt + \sigma dW)$. Si el tipo de interés continuo es r ya hemos visto que el precio hoy del contrato de futuros es

$$F_0 = S_0 e^{rT},$$

donde T es el tiempo hasta el vencimiento del futuro.

Queremos ahora describir el proceso (F_t) de precios del futuro. Con los mismos razonamientos con los que valoramos el futuro es fácil ver que, en cualquier instante $0 \leq t \leq T$ tenemos

$$F_t = S_t e^{r(T-t)} \tag{4}$$

Puesto que F es función $F(S,t)$, aplicamos el Lema de Itô como en (2) y tenemos

$$dF = \left(\frac{\partial F}{\partial S} \mu S + \frac{\partial F}{\partial t} + \frac{1}{2} \frac{\partial^2 F}{\partial S^2} \sigma^2 S^2 \right) dt + \frac{\partial F}{\partial S} \sigma S dW. \tag{5}$$

De (4) obtenemos el valor de las derivadas parciales que aparecen en 5 y llegamos a

$$dF = (e^{r(T-t)} \mu S - rSe^{r(T-t)})dt + e^{r(T-t)} \sigma S dW.$$

Usando ahora (4) llegamos a

$$dF = (\mu - r)Fdt + \sigma F dW,$$

es decir, el contrato de futuros sigue un proceso muy similar al del proceso de precios, con la única diferencia de la corrección en el drift.

16. Aparición de martingalas: Teorema de Girsanov

Estamos suponiendo que el proceso de precios del activo se puede escribir como

$$dS = S(\mu dt + \sigma dW)$$

o, equivalentemente

$$S_t = S_0 e^{(\mu - \frac{\sigma^2}{2})t + \sigma W_t}$$

Por tanto el proceso de precios descontados es

$$\tilde{S}_t = e^{-rt} S_t = S_0 e^{((\mu - r) - \frac{\sigma^2}{2})t + \sigma W_t}$$

y ya vimos cuando estudiamos cálculo estocástico que un proceso así es una martingala si y sólo si $\mu - r = 0$, es decir si $\mu = r$.

Por ello, para valorar pasaremos al proceso

$$dS = S(rdt + \sigma dW).$$

Lo que estamos haciendo es, en lugar de modificar la medida para que el proceso de precios descontados se convierta en una martingala, modificar el proceso de precios. No resulta totalmente obvio que ambas cosas sean equivalentes. Una demostración correcta de este hecho requiere del Teorema de Girsanov.

17. Valoración de una opción call

Veamos en esta sección cómo usar las técnicas de cálculo estocástico, junto con el Teorema Fundamental de Valoración de Activos, para valorar una opción call. Una vez entendido el proceso, es igualmente sencillo, salvo por la integración final, valorar otras clases de derivados.

Ya hemos visto que asumimos inicialmente que el proceso de precios es de la forma $dS = S(\mu dt + \sigma dW)$ pero que podemos suponer $\mu = r$, lo que convierte el proceso de precios descontados automáticamente en una martingala.

En ese caso, el Teorema Fundamental de Valoración de Activos nos dice que ahora valorar un derivado (de tipo europeo) no es más que integrar la función de pagos del derivado a vencimiento con respecto a la medida martingala y descontar ese valor a precios de hoy. Veamos cómo hacemos esto en el caso particular de una opción call.

La función de pagos de una call a vencimiento es

$$H(S,T) = (S_T - K)^+.$$

A vencimiento el precio es

$$S_T = S_0 e^{rT+\sigma W_T}$$

donde W_T está distribuido como una normal $N(0,\sqrt{T})$. Por tanto, el valor de la call viene dado por

$$C = e^{-rT}\mathbf{E}((S_T - K)^+) = e^{-rT}\frac{1}{\sqrt{2\pi}}\int_{\infty}^{\infty}(S_0 e^{(r-\frac{\sigma^2}{2})T+\sigma x\sqrt{T}} - K)^+ e^{-\frac{x^2}{2}}dx$$

Haciendo las cuentas llegamos a la conocida fórmula de Black-Scholes

$$C = e^{-rT}(S_0 e^{rT} N(d_1) - KN(d_2)) = S_0 N(d_1) - Ke^{-rT}N(d_2)),$$

donde

$$d_1 = \frac{\ln\frac{S_0}{K} + \left(r + \frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}},$$

y

$$d_2 = \frac{\ln\frac{S_0}{K} + \left(r - \frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}},$$

18. El enfoque por medio de EDP's: La ecuación diferencial de Black-Scholes

Hemos supuesto que el proceso de precios sigue una ecuación de la forma

$$dS = \mu S dt + \sigma S dW \tag{6}$$

Supongamos un derivado sobre S. Sea $f(S,t)$ el valor del derivado en función de S y t. (Aunque no lo necesitamos ahora, nótese que, en particular, $f(S,T)$ es la función de pagos del derivado a vencimiento en función del precio del subyacente.

Aplicando el Lema de Itô, tenemos que

$$df = \left(\frac{\partial f}{\partial S}\mu S + \frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial S^2}\sigma^2 S^2\right)dt + \frac{\partial f}{\partial S}\sigma S dW \tag{7}$$

Necesitamos ahora observar que la incertidumbre sobre el precio del derivado viene exclusivamente de la incertidumbre sobre el precio del activo, de manera que el Browniano que aparece en (6) y en (7) es el mismo.

Formamos entonces una cartera con -1 derivado y $\frac{\partial f}{\partial S}$ acciones. Entonces el valor de la cartera Π viene dado por

$$\Pi = -f + \frac{\partial f}{\partial S} S$$

y el cambio en el valor de la cartera en un intervalo Δt viene dado por

$$\Delta\Pi = -\Delta f + \frac{\partial f}{\partial S} \Delta S$$

Si sustituimos Δf y ΔS por sus valores dados por (7) y (6) tenemos

$$\Delta\Pi = \left(-\frac{\partial f}{\partial t} - \frac{1}{2} \frac{\partial^2 f}{\partial S^2} \sigma^2 S^2 \right) \Delta t$$

y observamos que en esta ecuación no aparece W, es decir, no es estocástica. Esto quiere decir que la cartera no tiene riesgo durante el intervalo Δt, de donde, por razonamientos de arbitraje, se debe tener

$$\Delta\Pi = r\Pi\Delta t$$

donde r es la tasa de interés libre de riesgo. Por tanto

$$\left(-\frac{\partial f}{\partial t} - \frac{1}{2} \frac{\partial^2 f}{\partial S^2} \sigma^2 S^2 \right) \Delta t = r \left(f - \frac{\partial f}{\partial S} S \right) \Delta t$$

Es decir

$$\frac{\partial f}{\partial t} + rS \frac{\partial f}{\partial S} + \frac{1}{2} \frac{\partial^2 f}{\partial S^2} \sigma^2 S^2 = rf$$

Esta es la ecuación de Black-Scholes.

Ahora, valorar el derivado es calcular $f(S_0, 0)$. Obsérvese que en la ecuación diferencial de Black-Scholes las condiciones de contorno no son iniciales sino finales, es decir conocemos $f(S, T)$. A partir de ese dato podemos implementar un esquema de diferencias finitas, o algún otro método de resolución numérico, para "resolver hacia atrás" y llegar a $f(S_0, 0)$.

19. Aparición de martingalas: versión continua 2. Razonamientos de arbitraje

Ya hemos visto anteriormente que para valorar un derivado sustituimos μ por r en el proceso de precios del subyacente. Hemos justificado este paso matemáticamente de dos maneras, una discretizando el proceso de precios y la otra utilizando el Teorema de Girsanov. Veamos ahora una tercera justificación esencialmente económica: En la ecuación diferencial de Black-Scholes no aparecen variables que involucren las preferencias de riesgo de los agentes (estas variables debieran ser o bien μ o bien, equivalentemente, la probabilidad que estamos asumiendo). Por tanto, si μ no aparece en la solución de la ecuación, podríamos haber asumido cualquier μ (equivalentemente cualquier medida) en la descripción del proceso de precios S. En particular, podríamos haber elegido $\mu = r$, y esto ya convierte automáticamente al proceso en una martingala.

20. Valoración mediante simulación Montecarlo

La idea a la hora de valorar mediante simulación es muestrear aleatoriamente nuestra fuente de incertidumbre con unas muestras distribuidas según el modelo que consideremos más adecuado, y promediar el valor que se obtendría para el derivado en cada uno de estos casos.

La fuente de incertidumbre que debemos muestrear depende del derivado. Si estamos valorando un derivado europeo, nos bastará con muestrear lo que sucede a vencimiento. Si estamos valorando un derivado *dependiente de la trayectoria* (*path dependent*) tendremos que simular la trayectoria entera seguida por el activo, discretizando el tiempo para considerar sólo un numero finito de instantes temporales.

El principal inconveniente de este método es que normalmente resulta computacionalmente muy costoso, además de que no resulta obvio cómo utilizarlo para valorar opciones americanas. Por otro lado, su principal ventaja es su versatilidad, puesto que permite elegir distintos modelos para el proceso S, y también tratar con facilidad los derivados dependientes de distintos activos.

Veamos cómo implementarlo. Supongamos como siempre hemos hecho hasta ahora que el proceso de precios viene dado por

$$dS = S(rdt + \sigma dW)$$

(trabajamos en el mundo neutral al riesgo). En ese caso, ya hemos visto en la Sección 3 que en ese caso tenemos

$$d \ln S = (\mu - \frac{\sigma^2}{2})dt + \sigma dW.$$

y por tanto

$$S_t = S_0 e^{(\mu - \frac{\sigma^2}{2})t + \sigma W_t}. \tag{8}$$

y

$$S_{t+\Delta t} = S_t e^{(\mu - \frac{\sigma^2}{2})\Delta t + \sigma Z \sqrt{\Delta t}}, \tag{9}$$

donde Z es una normal $N(0,1)$.

Cuando deseemos valorar un derivado en el que sólo nos interese lo que sucede a vencimiento, utilizaremos la fórmula (8). Cuando por el contrario tengamos que valorar un derivado dependiente del camino, utilizaremos la ecuación (9).

21. Valoración de otros derivados

21.1. Opciones binarias. Una opción de venta binaria, o digital, de tipo europeo con strike K, es una opción que paga 1 euro si el valor a vencimiento de la opción excede K, y paga 0 en caso contrario. Veamos cómo valorarla. De nuevo tendremos que calcular la integral de la función de pagos de la opción con respecto de la medida neutral al riesgo, y descontar ese valor para llegar a euros de hoy. Por tanto, llamando H a la función de Heaviside, el precio de una opción binaria debe ser

$$C_B = e^{-rT}\mathbf{E}(H(S_T - K)) = e^{-rT}P(S_T \geq K) =$$
$$= e^{-rT}\frac{1}{\sqrt{2\pi}}\int_{z_K}^{+\infty} e^{-\frac{x^2}{2}}dx = 1 - N(z_K) = N(-z_K),$$

donde N es la función de distribución de la normal de media 0 y varianza 1 y z_K es el único número que verifica

$$S_0 e^{(r-\frac{\sigma^2}{2})T + \sigma z_K\sqrt{T}} = K,$$

es decir,

$$z_K = \frac{1}{\sigma\sqrt{T}}\ln(\frac{Ke^{-rT}}{S_0}) + \frac{1}{2}\sigma\sqrt{T}.$$

Una opción binaria put es una opción que paga 1 si $S_T < K$ y 0 en caso contrario. Es trivial ver que si tenemos una put y una call binarias con el mismo strike, se tiene

$$P_B + C_B = e^{-rT},$$

de donde se sigue inmediatamente una expresión para el precio de la put P_B.

21.2. Opciones barrera. Una opción con strike K y barrera H es una opción (put o call) estándar que se activa (o se desactiva) si el valor del subyacente entre la fecha de emisión y la fecha de vencimiento supera (por arriba o por debajo) el valor H. Tenemos entonces opciones *up and out*, *up and in*, *down and out* y *down and in*, y para cada una de estas existen tanto opciones put como call. Por ejemplo, una opción put down and out con strike K y barrera H es una opción que paga $(K - S_T)^+$ euros a vencimiento, pero sólo en el caso de que en ningún momento entre las fechas de emisión y vencimiento el valor del subyacente haya descendido por debajo del valor H. Una opción put down and in con los mismos strikes y barrera es una opción que paga $(K - S_T)^+$ euros a vencimiento, pero sólo en el caso de que en algún momento entre las fechas de emisión y vencimiento el valor del subyacente haya descendido por debajo del valor H. Es claro que la suma de una put down and out mas una put down and in con los mismos strikes y vencimientos es una put estándar, de manera que sólo tenemos que valorar una de ellas. Lo mismo sucede con las restantes modalidades.

Veamos cómo podemos utilizar valoración neutral al riesgo para calcular el precio de estas opciones.

Vamos a necesitar algunos resultados técnicos previos.

Lema 21.1 (Principio de reflexión). *Sea (W_t) un movimiento browniano definido en $[0, T]$. Para todo $0 < B \leq A$ se tiene, para ΔB tendiendo a 0,*

$$P((\max_{0\leq t\leq T} W_t \geq A) \cap (W_T \in (B, B + \Delta B))) = \frac{1}{\sqrt{2\pi T}} e^{-\frac{(2A-B)^2}{2T}} \Delta B.$$

Demostración. Consideremos un camino ω que verifique la primera condición, es decir, que

$$\max_{0\leq t\leq T} W_t(\omega) \geq A.$$

Sea t_0 el primer instante t que verifica que $W_t(\omega) = A$ y consideramos el camino ω' "reflejado" con respecto de A a partir de t_0 de ω que verifica

$$W_t(\omega') = W_t(\omega) \text{ para todo } t \leq t_0$$

y

$$W_t(\omega') = 2A - W_t(\omega) \text{ para todo } t \geq t_0.$$

Es fácil convencerse de que los caminos ω y ω' son equiprobables (por supuesto cada uno de ellos tiene probabilidad 0, piénsese más bien en las versiones discretas, en paseos aleatorios, de estos caminos), de manera que

$$P((\max_{0\le t\le T} W_t \geq A) \cap (W_T \in (B, B+\Delta B))) =$$

$$= P((\max_{0\le t\le T} W_t \geq A) \cap (W_T \in (2A-B-\Delta B, 2A-B)))$$

puesto que $A < 2A-B$, se sigue que

$$P((\max_{0\le t\le T} W_t \geq A) \cap (W_T \in (B, B+\Delta B))) =$$

$$= P((\max_{0\le t\le T} W_t \geq A) \cap (W_T \in (2A-B-\Delta B, 2A-B))) =$$

$$= P(W_T \in (2A-B-\Delta B, 2A-B)) =$$

$$= \frac{1}{\sigma\sqrt{2\pi T}} \int_{2A-B-\Delta B}^{2A-B} e^{-\frac{x^2}{2T}} dx \sim \frac{1}{\sigma\sqrt{2\pi T}} e^{-\frac{(2A-B)^2}{2T}} \Delta B,$$

donde la última aproximación se convierte en una igualdad cuando ΔB tiene a 0. □

Usaremos la siguiente notación: El proceso $(W_t^\mu)_t$, al que denotaremos browniano con tendencia μ viene definido como

$$W_t^\mu = W_t + \mu t.$$

Necesitaremos una versión más fina de este resultado, en el que el Browniano se ve sustituido por un Browniano con drift. Para ello necesitamos previamente el siguiente resultado que enunciamos sin demostración. Una demostración puede verse en [**1**].

TEOREMA 21.2 (Cameron-Martin). *Let $F(w_1, w_2, \ldots, w_n)$ be a continuous function. Then*

$$\boldsymbol{E}(F(W_{t_1}^\mu, W_{t_2}^\mu, \ldots, W_{t_n}^\mu)) = \boldsymbol{E}(F(W_{t_1}, W_{t_2}, \ldots, W_{t_n}) e^{\mu W_{t_n} - \frac{1}{2}\mu^2 t_n}).$$

Con este resultado, ya podemos probar el siguiente

LEMA 21.3. *Sea (W_t^μ) un movimiento browniano con drift μ definido en $[0,T]$. Para todo $0 < B \leq A$ se tiene, para ΔB tendiendo a 0,*

$$P((\max_{0\le t\le T} W_t^\mu \geq A) \cap (W_T^\mu \in (B, B+\Delta B))) = \frac{1}{\sqrt{2\pi T}} e^{-\frac{(2A-B)^2}{2T}} e^{B\mu - \frac{1}{2}\mu^2 T} \Delta B.$$

Demostración. Consideramos la función

$$F(W_\mu) = \chi_{\left(\max_{0\le t\le T} W_t^\mu \ge A\right)\cap\left(W_T^\mu\in(B,B+\Delta B)\right)},$$

donde χ_H es la función característica del conjunto H. Utilizando la continuidad de (casi todos) los caminos del browniano, podemos aproximar la función F por una función evaluada en una cantidad finita de puntos a la que le podemos aplicar el Teorema 21.2, y llegamos a

$$\begin{aligned} P(\left(\max_{0\le t\le T} W_t^\mu \ge A\right) \cap (W_T^\mu \in (B, B+\Delta B))) = \\ = \mathbf{E}(\chi_{\left(\max_{0\le t\le T} W_t^\mu \ge A\right)\cap\left(W_T^\mu\in(B,B+\Delta B)\right)}) = \\ = \mathbf{E}(e^{\mu W_T - \frac{1}{2}\mu^2 T}\chi_{\left(\max_{0\le t\le T} W_t \ge A\right)\cap\left(W_T^\mu\in(B,B+\Delta B)\right)}) = \\ = \frac{1}{\sqrt{2\pi T}} e^{-\frac{(2A-B)^2}{2T}} e^{B\mu-\frac{1}{2}\mu^2 T}\Delta B, \end{aligned}$$

donde la última igualdad se sigue del Lema 21.1. □

Con estos resultados a nuestra disposición ya podemos valorar una put down and out con barrera.

Suponemos como siempre que el proceso de precios, en el mundo neutral al riesgo, es

$$dS = S(rdt + \sigma dW).$$

Usando que el precio es el valor descontado de la integral con respecto de la medida neutral al riesgo del pago del derivado tenemos, llamando P_{do} al precio de la put down and out con strike K y barrera H, que

$$\begin{aligned} P_{do} = e^{-rT}\mathbf{E}((K-S_T)^+\chi_{\min_{0\le t\le T} S_t > H}) = \\ = e^{-rT}\mathbf{E}((K-S_T)\cdot\chi_{\min_{0\le t\le T} S_t > H}\cdot\chi_{H<S_T<K}) = \\ = e^{-rT}\mathbf{E}((K-S_T)\cdot\chi_{H<S_T<K}) - \\ -e^{-rT}\mathbf{E}((K-S_T)\cdot\chi_{\min_{0\le t\le T} S_t \le H}\cdot\chi_{H<S_T<K}) = \\ = e^{-rT}\mathbf{E}((K-S_T)\cdot\chi_{S_T<K}) - e^{-rT}\mathbf{E}((K-S_T)\cdot\chi_{S_T<H}) - \\ -e^{-rT}KP((\min_{0\le t\le T} S_t \le H)\cap(H<S_T<K)) + \\ +e^{-rT}\mathbf{E}(S_T\cdot\chi_{\min_{0\le t\le T} S_t\le H}\cdot\chi_{H<S_T<K}). \end{aligned}$$

Fijémonos en los cuatro sumandos de la última igualdad. El primero es una put estándar con strike K. El segundo se puede valorar fácilmente usando razonamientos análogos a los utilizados para la valoración de una opción call (o put) estándar.

Veamos como calcular los dos últimos términos. Definimos las siguientes cantidades

$$A_H = \frac{1}{\sigma} \ln \frac{H}{S_0}$$
$$A_K = \frac{1}{\sigma} \ln \frac{K}{S_0}$$
$$\mu = \frac{r}{\sigma} - \frac{1}{2}\sigma.$$

Con esta notación se puede ver que $H < S_T$ si y sólo si

$$A_H = \frac{1}{\sigma} \ln \frac{H}{S_0} < \frac{1}{\sigma} \ln \frac{S_T}{S_0} = \frac{1}{\sigma}((r - \frac{\sigma^2}{2})t + \sigma W_t) =$$
$$= (\frac{r}{\sigma} - \frac{\sigma}{2})t + W_t = W_t^\mu$$

y por tanto, usando los procesos $-W_t^\mu = -W_t - \mu t$ y $-W_t^{-\mu} = W_t - \mu t$ tienen la misma distribución de probabilidad, por la simetría del browniano, tenemos

$$P((\min_{0 \leq t \leq T} S_t \leq H) \cap (H < S_T < K)) =$$
$$= P((A_H < W_T^\mu < A_K) \cap (\min_{0 \leq t \leq T} W_t^\mu \leq A_H)) =$$
$$= P((-A_K < W_T^{-\mu} < -A_H) \cap (\max_{0 \leq t \leq T} W_t^{-\mu} \geq -A_H)) =$$
$$= \int_{-A_K}^{-A_H} \frac{1}{\sqrt{2\pi T}} e^{-\frac{(2A_H - B)^2}{2T}} e^{B\mu - \frac{1}{2}\mu^2 T} dB.$$

Análogamente, para el cuarto término, tenemos

$$\mathbf{E}(S_T \cdot \chi_{\min_{0 \leq t \leq T} S_t \leq H} \cdot \chi_{H < S_T < K}) =$$
$$= S \int_{-A_K}^{-A_H} \frac{1}{\sqrt{2\pi T}} e^{-\sigma B} e^{-\frac{(2A_H + B)^2}{2T}} e^{-B\mu - \frac{1}{2}\mu^2 T} dB.$$

Las dos integrales se pueden calcular explícitamente, y los dos términos iniciales se pueden valorar conociendo los valores de call y put estándar. Juntando todo esto, se tiene

$$P_{do} = Ke^{-rT} N(-d_2^k) - SN(-d_1^K) - Ke^{-rT} N(-d_2^H) + SN(-d_1^H) -$$
$$-Ke^{-rT} \left(\frac{H}{S}\right)^{\left(\frac{2r}{\sigma^2} - 1\right)} (N(d_4) - N(d_5)) + S\left(\frac{H}{S}\right)^{\left(\frac{2r}{\sigma^2} + 1\right)} (N(d_6) - N(d_7)),$$

donde

$$d_1^K = \frac{1}{\sigma\sqrt{T}} \left(\ln\left(\frac{S}{K}\right) + \left(r + \frac{\sigma^2}{2}\right) T \right),$$

$$d_2^K = \frac{1}{\sigma\sqrt{T}}\left(\ln\left(\frac{S}{K}\right) + \left(r - \frac{\sigma^2}{2}\right)T\right),$$

$$d_1^H = \frac{1}{\sigma\sqrt{T}}\left(\ln\left(\frac{S}{H}\right) + \left(r + \frac{\sigma^2}{2}\right)T\right),$$

$$d_2^H = \frac{1}{\sigma\sqrt{T}}\left(\ln\left(\frac{S}{H}\right) + \left(r - \frac{\sigma^2}{2}\right)T\right),$$

$$d_4 = \frac{1}{\sigma\sqrt{T}}\left(\ln\left(\frac{H}{S}\right) + \left(r - \frac{\sigma^2}{2}\right)T\right),$$

$$d_5 = \frac{1}{\sigma\sqrt{T}}\left(\ln\left(\frac{H^2}{SK}\right) + \left(r - \frac{\sigma^2}{2}\right)T\right),$$

$$d_6 = \frac{1}{\sigma\sqrt{T}}\left(\ln\left(\frac{H}{S}\right) + \left(r + \frac{\sigma^2}{2}\right)T\right),$$

$$d_7 = \frac{1}{\sigma\sqrt{T}}\left(\ln\left(\frac{H^2}{SK}\right) + \left(r + \frac{\sigma^2}{2}\right)T\right),$$

22. Acciones que pagan dividendos

Supongamos ahora que el activo tiene una rentabilidad continua por dividendos q anual. Este sería el caso de una divisa, el caso de acciones que pagan dividendos puntuales siempre podemos tratarlo convirtiéndolos en un dividendo continuo equivalente.

Hay varias formas posibles de razonar. Utilicemos la noción de medida neutral al riesgo. Ya hemos visto anteriormente que en el mundo neutral al riesgo, el rendimiento (medio) de cualquier activo debe ser el tipo de interés r. Puesto que nuestro activo ya tiene un rendimiento q, su deriva esperada debe ser $\mu = r - q$ para que el proceso de precios descontados sea una martingala. Por tanto, debemos razonar cómo lo hemos venido haciendo hasta ahora, utilizando como proceso para el activo

$$dS = S((r-q)dt + \sigma dW_t)$$

en lugar de

$$dS = S(rdt + \sigma dW_t).$$

Podemos detallar esto de la siguiente manera. Suponemos que tenemos inicialmente 1 unidad del activo, y suponemos que reinvertimos su rendimiento (continuo) en el propio activo. En ese caso tendremos en cada momento e^{qt} unidades del activo, valoradas en $S_t e^{qt}$ euros. Si suponemos que el proceso de precios S_t viene dado por

$$S_t = S_0 e^{\left((r-q)-\frac{\sigma^2}{2}\right)t+\sigma W_t},$$

entonces el valor de nuestra inversiï£¡n en el instante t serï£¡

$$S_t e^{qt} = S_0 e^{\left(r-\frac{\sigma^2}{2}\right)t+\sigma W_t},$$

tal y como debe ocurrir para que el proceso de precios descontados sea una martingala.

Observad que los valores que se obtienen al valorar de esta manera son los mismos que se obtendrían si valoráramos un derivado que no paga dividendos con precio inicial $S_0 e^{-qT}$. Esta fórmula se podría obtener del siguiente razonamiento alternativo: cuando una acción paga un dividendo, su valor decrece exactamente en la cantidad pagada. Podemos entonces descomponer el valor de la acción en una componente determinista, que decrece exponencialmente con e^{-qt} y una componente estocástica, de valor inicial $S_0 e^{-qT}$

Si buscamos implementar estos razonamientos en un árbol, de nuevo por razonamientos de arbitraje, basta con razonar como lo hacíamos anteriormente eligiendo ahora la medida neutral al riesgo de manera que el crecimiento medio ya no sea $e^{r\Delta t}$ sino $e^{(r-q)\Delta t}$.

En ciertos casos, no podemos considerar equivalente un (o varios) único pago de dividendos en un momento puntual a un pago continuo (considerad por ejemplo la valoración de opciones barrera). En el árbol binomial podemos estudiar ese caso de la siguiente forma. Supongamos que sabemos que en el instante i se va a producir un pago de dividendos por una cantidad $\alpha_i S_i$, donde S_i es el valor del activo en el instante i. En ese caso, el valor en el instante S_{i+1} sera $S_i(1-\alpha_i)u_i$ o bien $S_i(1-\alpha_i)d_i$, donde como siempre $u_i, d_i = e^{\pm\sigma\sqrt{\Delta t}}$

23. Tipos de interés variables

Es fácil introducir tipos de interés variables en los árboles binomiales. Recordad que hacíamos los razonamientos sobre el proceso de precios descontados. Hasta ahora habíamos supuesto que el factor de descuento era constante en el tiempo, pero en realidad esa suposición no se usa en ninguno de los razonamientos hechos en el árbol binomial, y podríamos perfectamente descontar con tipos distintos en cada uno de los periodos.

En simulación Montecarlo también es obvio cómo introducir tipos de interés variables. Simplemente tenemos que utilizar el tipo de interés adecuado en cada paso de la simulación del camino.

24. Volatilidades variables

No es obvio cómo introducir volatilidades variables en un árbol binomial o trinomial. Por ello, no hablaremos de ello ahora.

Por el contrario, de nuevo a la hora de simular Montecarlo es muy fácil utilizar volatilidades: en cada paso de la simulación utilizamos la volatilidad correspondiente a ese paso.

25. Valoración de opciones barrera con árboles

26. Árboles trinomiales

Veamos ahora cómo generar un árbol trinomial. La idea es exactamente la misma que en el caso de árboles binomiales, salvo que ahora en cada nodo tenemos tres posibilidades en lugar de dos. Si en el nodo i-ésimo el valor es S, entonces en el nodo $i+1$-ésimo los posibles valores serán $SU > SM > SD$. Querremos de nuevo que el árbol recombine, para que el número de nodos no crezca exponencialmente. Es fácil ver que una condición necesaria y suficiente para que esto suceda es que

$$UD = M^2.$$

Tenemos que elegir U, D, M y asignar probabilidades P_U, P_M, P_D a los respectivos sucesos, de manera que el proceso de precios descontados sea una martingala respecto de esta probabilidad, y de manera que la media y la varianza del proceso descrito por el ï£¡rbol coincidan con la media y la varianza del proceso

$$dS = S(rdt + \sigma dW_t)$$

que queremos aproximar.

Hay infinitas formas de hacer esto. Las elecciones que aparecen en **[8]** son

$$U = e^{\sigma\sqrt{3\Delta t}}, \quad M = 1, \quad D = \frac{1}{U}, \quad P_U = \sqrt{\frac{\Delta t}{12\sigma^2}}\left(r - \frac{\sigma^2}{2}\right) + \frac{1}{6},$$

$$P_M = \frac{2}{3}, \quad P_D = -\sqrt{\frac{\Delta t}{12\sigma^2}}\left(r - \frac{\sigma^2}{2}\right) + \frac{1}{6}$$

Capítulo 6

Gestión de Riesgos Financieros

1. Introducción

Probablemente la palabra clave en gestión financiera sea *riesgo*. A la hora de manejar cualquier producto financiero es fundamental entender el riesgo que este comporta. Es necesario cuantificar dicho riesgo y, muy a menudo, cubrirse total o parcialmente de él.

Existen muchas fuentes de riesgo, y por ello se suele dividir su estudio en distintas categorías. Entre ellas

- Riesgo de Mercado
- Riesgo de Crédito
- Riesgo de Liquidez
- ...

1.1. Riesgo de Mercado. El *riesgo de mercado* es el asociado a la variación del valor de mercado de un producto. Esencialmente es el riesgo de que un producto que poseemos baje de valor. Es en este en el que nos concentraremos en este módulo.

1.2. Riesgo de Crédito. El *riesgo de crédito* es el asociado a que un agente que nos tenga que hacer unos pagos incumpla su compromiso (voluntariamente o por quiebra).

1.3. Riesgo de liquidez. El *riesgo de liquidez* es el asociado a que un instrumento que debamos vender (o comprar) en cierto momento no sea líquido, y por tanto no podamos ejecutar la operación.

2. Riesgo de Mercado

Como ya hemos dicho antes, el Riesgo de Mercado es el asociado a la variación del valor de mercado de los activos que poseamos. La situación es siempre esencialmente la siguiente:

Tenemos una cartera formada por una serie de activos financieros. A fecha de hoy conocemos el valor de la cartera; en cambio, este valor en cualquier instante futuro no es conocido, y lo modelizaremos mediante una variable aleatoria definida sobre un espacio de probabilidad del que sabemos muy poco, el espacio de los “estados del mundo”. Nuestro objetivo es poder decir algo acerca de la función de distribución de esta variable aleatoria. En concreto, cuando estamos estudiando el *riesgo* de nuestra cartera, nuestro objetivo es poder decir algo acerca de "cómo de mal" podrían ir las cosas.

De una manera muy general, la manera de realizar este estudio es

1. Identificar las fuentes de incertidumbre que afecten al valor de nuestra cartera.
2. Estudiar la variación de estas fuentes de incertidumbre.
3. Estudiar cómo dicha variación afecta al valor de nuestra cartera.

A la hora de efectuar cualquiera de los pasos anteriores existen varias posibilidades.

En el primer paso, podemos elegir un gran nivel de detalle, identificando y tratando por separado todas las fuentes de riesgo, pero en ocasiones esto no es posible y resulta más razonable quedarse sólo con las fuentes más importantes de incertidumbre, o utilizar técnicas de *componentes principales*, o análogas, para seleccionar un número reducido de variables que sin embargo aporten suficiente información.

En el segundo paso, a la hora de estudiar cómo varían estas fuentes de incertidumbre, existen dos metodologías principales. Una de ellas consiste en examinar la serie histórica de la variable, y a partir de ahí deducir información para el futuro. La otra consiste en aceptar un *modelo* para la variable y extraer de él la información. Ambos métodos están relacionados, puesto que un modelo sólo es aceptado si ajusta suficientemente bien series históricas anteriores. Más adelante utilizaremos ambos métodos para analizar riesgos.

Para el tercer paso, cuando el valor de la cartera está relacionada mediante una fórmula sencilla con el valor de las variables, no hay problema. Pero cuando está relación es más compleja, en ocasiones recurriremos a métodos de aproximación. Veremos ejemplos de esto más adelante.

3. Midiendo el Riesgo

Suponemos que tenemos una cartera de valores. Su valor hoy es un dato conocido y su valor en un determinado instante temporal futuro es una variable aleatoria. A partir de esta, podemos calcular la variable aleatoria que mide la *variación de valor* de la cartera. Entonces, *medir el riesgo* de la cartera consiste en considerar la función de distribución de la variación de valor y calcular uno o varios estadísticos asociados a valores negativos de dicha variación. Concretemos:

La metodología habitual es comenzar calculando el riesgo a 1 día. Tenemos una cartera cuyo valor hoy es V_0. Su valor mañana es una variable aleatoria V_1. A partir de V_0 y V_1 podemos calcular el *rendimiento* a 1 día de la cartera.

Al igual que cuando veíamos tipos de interés, podemos pensar en rendimientos simples, compuestos o continuos. Cuando pensemos en el rendimiento *simple* a 1 día, lo definiremos como

$$R = \frac{V_1 - V_0}{V_0}$$

mientras que si pensamos en rendimiento *continuo* a 1 día lo definimos como

$$R = \ln \frac{V_1}{V_0}.$$

Cada una de las definiciones anteriores tiene ventajas e inconvenientes, y en cada ocasión elegiremos la forma de trabajar que resulte más conveniente.

Una vez que tenemos definido el rendimiento a 1 día R, veamos qué estadísticos de R nos interesa estudiar para medir el riesgo de nuestra cartera.

El más utilizado es VaR (*Value at Risk*). Es el estadístico necesario para poder hacer una afirmación del tipo “Estamos X por ciento seguros de que nuestras perdidas a un día no excederán del Y del valor de nuestra cartera”.

Es decir, se define VaR a un día con un umbral de confianza X como el percentil $100 - X$ de la variable R. Prácticamente siempre, dicho percentil serán perdidas. Como norma lo cambiamos de signo y lo expresamos como una cantidad positiva.

Por ejemplo, decir que el VaR99 % es el 2 % significa que estamos 99 % seguros de que las perdidas no excederán del 2 %, equivalentemente que el percentil 1 % es 2 %.

Formalmente, si llamamos VaR_α a VaRα % se tiene

$$P(R < -\mathrm{VaR}_\alpha) = (100 - \alpha)\,\%$$

Obviamente VaR es una medida muy simple, y eso es una ventaja y un inconveniente. La principal ventaja es que se sigue del hecho de que es fácil de entender. Por ello es el principal medidor de riesgo usado por los gestores y por los organismos reguladores. El principal inconveniente es que pierde mucha información. En particular, las dos siguientes distribuciones de probabilidad de los rendimientos tienen el mismo VaR, mientras que la segunda es claramente mucho más arriesgada que la primera.

Por este motivo, se han definido también otras medidas del riesgo que, aunque menos usadas, pretenden compensar este fallo de VaR. La más significativa es TailVar.

TailVaR responde a la pregunta “Si las cosas van mal, ¿cómo de mal irán?”. TailVaR$\alpha\,\%$ es la perdida media *condicionado a que haya perdidas* superiores al nivel marcado por el percentil $100-\alpha$. Es decir

$$\mathrm{TailVaR}_\alpha = -E(R|R < -\mathrm{VaR}_\alpha)$$

3.1. Cambio de horizonte temporal. Como ya hemos dicho, la metodología habitual es calcular el VaR a un día y a partir de él calcular el VaR para periodos más largos de tiempo. Si queremos calcular el VaR a N días, la fórmula es

$$\text{VaR a } N \text{ días} = \text{VaR a 1 día} \times \sqrt{N}$$

Justificaremos esta fórmula más adelante, cuando veamos que los modelos para las cotizaciones de acciones asumen que las variaciones de valor son (esencialmente) normales independientes idénticamente distribuidas.

4. Simulación histórica

Veremos en esta sección cómo se usan datos históricos de cotizaciones para calcular el VaR a un día.

Supondremos a lo largo de esta sección que tenemos una cartera formada por 20 acciones de la compañía X, 30 acciones de la compañía Y y 4 opciones de compra sobre acciones de la compañía X con un precio de ejercicio de 80 y vencimiento dentro de 6 meses.

Conocemos la fórmula para valorar las opciones, que viene dada por

$$C = S\Phi(d_1) - Ke^{-rT}\Phi(d_2)$$

donde C es el precio de la opción, S la cotización actual de X, K es el precio de ejercicio, r es el tipo de interés (a un año), T es el tiempo hasta vencimiento (medido en años), σ es la volatilidad de X, Φ es la función de distribución de la normal estándar,

$$d_1 = \frac{\ln\frac{S}{K} + (r + \frac{\sigma^2}{2})T}{\sigma\sqrt{T}}$$

y

$$d_2 = d_1 - \sigma\sqrt{T}$$

o bien

$$d_\pm = \frac{\ln\frac{S}{K} + (r \pm \frac{\sigma^2}{2})T}{\sigma\sqrt{T}}$$

con $d_1 = d_+$ y $d_2 = d_-$

Suponemos también que conocemos los valores de σ, r y las cotizaciones de X e Y durante los últimos N días, con N un número grande, del orden de 1000 días.

Veamos cómo usamos la serie histórica. Sólo nos preocupa la variación de 1 día al siguiente, puesto que estamos estudiando el VaR a 1 día. Por ello consideramos las variaciones que ha habido en el valor de X, Y, σ y r desde cada uno de los días al siguiente. Obtenemos así $N-1$ valores históricos reales de cómo ha sido la variación de esos valores en cada uno de esos días. Con cada uno de esos datos obtenemos un posible *escenario* para mañana: los valores de hoy perturbados por la correspondiente variación histórica. Suponemos que cada uno de esos escenarios son equiprobables y que en conjunto son capaces de dibujar un contorno realista de la función de distribución de los valores de mañana. Entonces para cada uno de esos valores calculamos el valor de la cartera y a partir de ahí calculamos los estadísticos

EJERCICIO 4.1. *Calcular el VaR y TailVaR de una cartera usando datos históricos.*

Observación 4.2. *Cuando subimos el umbral para el cálculo del VaR, nos concentramos en la búsqueda de eventos extremos. En ese caso, siempre nos tendremos que enfrentar al problema de que no dispondremos de suficientes datos. En ocasiones es conveniente usar resultados de* Extreme Value Theory *para estudiar esos problemas*

5. Simulación mediante modelo

La alternativa a la simulación histórica es la simulación mediante modelo. En este caso elegimos un modelo que nos describa lo que creemos que van a ser los precios de nuestros activos. Ambos enfoques están más relacionados de lo que puede parecer a primera vista: en la simulación histórica asumimos que las variaciones de un día al siguiente son independientes entre sí, lo que pertenece a un modelo. En el otro sentido, cuando asumimos que un modelo es cierto, lo calibramos mediante la observación de datos históricos.

6. El modelo normal

Al igual que hacíamos a la hora de valorar derivados, asumiremos habitualmente el *modelo normal* para la cotización de una activo. Es decir

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t,$$

lo que, usando el Lema de Ito, nos da

$$S_t = S_0 e^{(\mu - \frac{\sigma^2}{2})t + \sigma W_t}.$$

A los efectos de la medición de riesgo, el cambio en el valor provocado por la tendencia μ es despreciable frente al provocado por la volatilidad σ (y además $\sigma^2 << \sigma$) de manera que en nuestros modelo suponemos que $\mu - \frac{\sigma^2}{2} = 0$

Asumimos por tanto que el rendimiento diario de una acción vendrá descrito por una normal de media 0 y desviación típica σ

$$\ln \frac{V_1}{V_0} = \sigma Z.$$

Nos referiremos habitualmente a σ como la *volatilidad*. Típicamente la volatilidad viene expresada en términos anuales. Aunque está implícito en el modelo, veamos ahora explícitamente cómo recuperar a partir de la volatilidad anual la volatilidad diaria.

Sea σ_1 la volatilidad a 1 día y σ_2 la volatilidad a 2 días. Es decir, tenemos que

$$\ln\frac{V_1}{V_0}=\sigma_1 Z,$$
$$\ln\frac{V_2}{V_1}=\sigma_1 Z,$$
$$\ln\frac{V_2}{V_0}=\sigma_2 Z,$$

donde Z es una variable aleatoria distribuida como una normal $N(0,1)$.

Entonces

$$\sigma_2 Z=\ln\frac{V_2}{V_0}=\ln\frac{V_2V_1}{V_0V_1}=\ln\frac{V_2}{V_1}+\ln\frac{V_1}{V_0}=\sigma_1 Z+\sigma_1 Z$$

Puesto que asumimos que los saltos son independientes entre sí, tenemos que el término de la derecha es la suma de 2 normales independientes y sabemos que esto es una normal de media cero (la suma de las medias) y desviación típica $\sigma_1\sqrt{2}$ (la varianza es la suma de las varianzas). De ahí se sigue que $\sigma_2=\sqrt{2}\sigma_1$. Aplicando los mismos razonamientos obtenemos que la volatilidad a un año es $\sqrt{252}$ por la volatilidad a un día (recordemos que hay 252 días de mercado al año).

6.1. Un único activo. Una vez que asumimos que los rendimientos vienen descritos por una normal, resulta muy fácil calcular el VaR de un único activo:

Tenemos que

$$P(\sigma Z<-\mathrm{VaR}_\alpha)=1-\alpha$$

por tanto

$$P(Z<\frac{-\mathrm{VaR}_\alpha}{\sigma})=1-\alpha$$

y, llamando Φ a la función de distribución de la normal $N(0,1)$, se tiene que

$$\Phi^{-1}(1-\alpha)=-\frac{\mathrm{VaR}_\alpha}{\sigma}$$

Podemos calcular también una fórmula cerrada para el TailVaR. Como

$$\mathbf{E}(Z|Z<t)=\frac{\mathbf{E}(Z\chi_{Z<t})}{P(Z<t)}=\frac{1}{P(Z<t)}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{t}ze^{-\frac{z^2}{2}}dz=\frac{1}{P(Z<t)\sqrt{2\pi}}e^{-\frac{t^2}{2}}$$

se tiene por tanto que

$$\mathrm{TailVaR}_\alpha=-\mathbf{E}(\sigma Z|\sigma Z<-\mathrm{VaR}_\alpha)=-\sigma\mathbf{E}(Z|Z<\frac{-\mathrm{VaR}_\alpha}{\sigma})=$$

$$= \frac{\sigma}{P(Z < \frac{-VaR_\alpha}{\sigma})\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{VaR_\alpha}{\sigma}\right)^2}$$

Puesto que

$$P(Z < \frac{-VaR_\alpha}{\sigma}) = 1 - \alpha$$

y que

$$\frac{-\mathrm{VaR}_\alpha}{\sigma} = \Phi^{-1}(1-\alpha)$$

llegamos finalmente a

$$\mathrm{TailVaR}_\alpha = \frac{\sigma}{(1-\alpha)\sqrt{2\pi}} e^{-\frac{(\Phi^{-1}(1-\alpha))^2}{2}}$$

6.2. Varios activos. Todo lo anterior describe el comportamiento de una acción. La situación varía al considerar varias acciones. En ese caso, los saltos finitos de cada una de ellas se modelizan mediante normales, pero estas normales *no* son independientes entre sí, sino que guardan una cierta correlación. Veremos cómo incluir el efecto de las correlaciones en nuestro modelo.

Inicialmente consideramos una cartera con n activos. El valor de cada activo hoy es V_0^i y el valor de la cartera es $V_0^c = \sum_i V_0^i$. Denotamos por p_i el peso que el activo i tiene hoy en la cartera. Se tiene que

$$p_i = \frac{V_0^i}{V_0^c}.$$

Trabajando con rendimientos simples (o aproximando los rendimientos compuestos por Taylor linealmente) es fácil ver que el rendimiento de la cartera es la media ponderada (con los pesos p_i) de los rendimientos de cada activo por separado:

$$R_c = \frac{\sum_i V_1^i - \sum_i V_0^i}{\sum_i V_0^i} = \sum_i \frac{V_1^i - V_0^i}{V_0^c} = \sum_i \frac{V_1^i - V_0^i}{V_0^i}\frac{V_0^i}{V_0^c} = \sum_i p_i \frac{V_1^i - V_0^i}{V_0^i}$$

El rendimiento de cada activo $\ln \frac{V_1^i}{V_0^i}$ está distribuido como $\sigma_i Z_i$ donde Z_i es una normal de media 0 y desviación típica 1.

Lo relevante aquí es que las variables $Z_1, \ldots, Z_n$ no son independientes.

Lo que hacemos entonces es considerar $(Z_1, \ldots, Z_n)$ como una distribución multidimensional. Asociada a ella tenemos la matriz

$$\Omega = \begin{pmatrix} \rho_{11} & \rho_{12} & \cdots & \rho_{1n} \\ \rho_{21} & \rho_{22} & \cdots & \rho_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ \rho_{n1} & \rho_{n2} & \cdots & \rho_{nn} \end{pmatrix}$$

en la que ρ_{ij} es el coeficiente de correlación entre Z_i y Z_j.

Ya hemos visto que el rendimiento de la cartera es la media ponderada con los pesos p_i de las rendimientos de cada activo, es decir

$$R^c = \sum_{i=1}^{n} \pi_i \sigma_i Z_i.$$

La media va a ser 0 y la Teoría de la Probabilidad nos dice que su varianza vale

$$\begin{pmatrix} p_1 & p_2 & \cdots & p_n \end{pmatrix} \begin{pmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \sigma_n \end{pmatrix} \Omega \begin{pmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \sigma_n \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ \cdots \\ p_n \end{pmatrix}$$

equivalentemente

$$\begin{pmatrix} p_1\sigma_1 & p_2\sigma_2 & \cdots & p_n\sigma_n \end{pmatrix} \Omega \begin{pmatrix} p_1\sigma_1 \\ p_2\sigma_2 \\ \cdots \\ p_n\sigma_n \end{pmatrix}$$

de manera que la desviación típica de los rendimientos de la cartera viene dada por

$$\sigma^c = \left(\sum_{i,j} p_i p_j \sigma_i \sigma_j \rho_{i,j} \right)^{\frac{1}{2}}$$

y con este dato ya podemos calcular los estadísticos.

6.3. La diversificáción reduce el riesgo. Supongamos dos activos A y B cada uno de ellos con la misma volatilidad $\sigma_A = \sigma_B = \sigma$ y con una constante de correlación entre ellos ρ. Si tuviéramos una cartera con uno cualquiera de los dos activos, su volatilidad sería σ. Si en cambio formamos una cartera con proporciones de A y B dadas por p_A y $p_B = 1 - p_A$, entonces según la sección anterior su volatilidad sería

$$\left(\begin{pmatrix} p_A\sigma & p_B\sigma \end{pmatrix} \times \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \times \begin{pmatrix} p_A\sigma \\ p_B\sigma \end{pmatrix} \right)^{\frac{1}{2}} = \left(\sigma^2 (p_A^2 + p_B^2 + 2 p_A p_B \rho) \right)^{\frac{1}{2}} =$$

$$= \left(\sigma^2(p_A^2 + p_B^2 + 2p_Ap_B - 2p_Ap_B(1-\rho))\right)^{\frac{1}{2}} =$$
$$= \left(\sigma^2((p_A + p_B)^2 - 2p_Ap_B(1-\rho))\right)^{\frac{1}{2}} = \left(\sigma^2(1 - 2p_Ap_B(1-\rho))\right)^{\frac{1}{2}} \leq \sigma$$

y el igual en la desigualdad anterior se obtiene únicamente si $\rho = 1$, es decir, si ambos activos son esencialmente el mismo.

6.4. Derivados. Cuando calculamos el riesgo asociado a una cartera de acciones, la variación en el valor de cada componente de la cartera depende linealmente de la variación del precio de la acción respectiva. Eso simplifica las cuentas que hemos tenido que realizar hasta el momento.

Supongamos ahora que introducimos derivados en nuestra cartera. Puesto que no tenemos un modelo que nos dé directamente la evolución de los precios del derivado, tenemos que simular la evolución de los precios del subyacente (y quizás otros factores como la volatilidad, tipos de interés, etc) y a partir de ellos calcular el riesgo del derivado. El problema que tendremos en ese caso es que la variación de los precios del derivado no es lineal con respecto a ninguno de esos factores. Esto imposibilita la obtención de fórmulas cerradas para el VaR o el TailVaR.

Tenemos entonces varias posibilidades. Una de ellas es considerar aproximaciones lineales (Taylor de orden 1) y tratar el problema de manera similar a cómo lo hemos hecho hasta ahora. El problema con esto es que la aproximación de orden 1 es demasiado pobre en general.

Por ello, otra posibilidad es considerar aproximaciones cuadráticas (Taylor de orden 2). Describimos a continuación cómo hacer esto.

Tenemos un derivado cuyo valor, que denotamos por C, depende de un subyacente X. Típicamente el valor del derivado depende de varios factores (cotización de X, volatilidad de X, tipo de interés, tiempo a vencimiento, etc). Inicialmente sólo consideraremos el riesgo inducido por la variación de valor de X, de manera que consideramos los demás factores constantes y escribimos

$$C = f(X).$$

En general f no es lineal, y por ello utilizamos su polinomio de Taylor.

Se denota por *Delta*, Δ, la sensibilidad de C a X, es decir

$$\Delta = \frac{dC}{dX}$$

y se denota por *Gamma*, Γ, la sensibilidad de la Delta al subyacente, es decir

$$\Gamma = \frac{d\Delta}{dX} = \frac{d^2C}{dX^2}.$$

EJERCICIO 6.1. *Sabiendo que el precio de una opción C viene dado por la fórmula de Black-Scholes descrito anteriormente, verificar que*

$$\Delta = \Phi(d_1)$$

y

$$\Gamma = \phi\left(\frac{d_1}{S\sigma\sqrt{T}}\right)$$

donde $\phi = \Phi'$ es la función de densidad de la normal estándar.

Consideremos inicialmente la aproximación lineal, es decir tenemos en cuenta solo Δ. En ese caso aproximamos

$$C_1 - C_0 \sim \Delta(X_1 - X_0)$$

y tenemos

$$\frac{C_1 - C_0}{C_0} = \frac{\Delta(X_1 - X_0)}{C_0} = \Delta\frac{X_0}{C_0}\frac{(X_1 - X_0)}{X_0}.$$

Usamos ahora que nuestro modelo asume rendimientos de X normales de la forma $\sigma_X Z$ y llegamos a

$$\frac{C_1 - C_0}{C_0} = \Delta\frac{X_0}{C_0}\sigma_X Z$$

Podemos llamar

$$\sigma_C = \Delta\frac{X_0}{C_0}\sigma_X$$

y finalmente tratamos los rendimientos del derivado de manera similar a como tratamos los rendimientos del activo.

EJERCICIO 6.2. *Implementar lo anterior en Excel, calculando el riesgo asociado a una cartera que consiste de 2 acciones y un derivado sobre una de ellas, usando la aproximación lineal.*

6.5. Aproximación cuadrática. Como ya hemos dicho, la aproximación lineal puede resultar a menudo demasiado pobre. En ese caso pasamos a considerar también la Γ y aproximamos la variación del valor del derivado como

$$C_1 - C_0 \sim \Delta(X_1 - X_0) + \frac{\Gamma}{2}(X_1 - X_0)^2.$$

Suponiendo de nuevo rendimientos normales para X llegamos a que

$$\frac{C_1 - C_0}{C_0} = \left(\frac{X_0}{C_0}\sigma_X\Delta\right) Z + \left(\frac{X_0^2}{C_0}\frac{\Gamma}{2}\sigma_X^2\right) Z^2$$

Como podemos ver esta distribución ya no es normal, puesto que es la suma de una normal más el cuadrado de una normal. Eso hará que no podamos obtener fórmulas cerradas para el cálculo del VaR de un derivado cuando usemos la aproximación cuadrática.

Por ello, a la hora de calcular los estadísticos de riesgo (VaR, TailVaR o cualquier otro) nos vemos obligados a hacerlo mediante simulación Montecarlo.

Describimos aquí los pasos que hemos de dar para realizar esto.

1. En primer lugar simulamos tantas variables (habitualmente normales, aunque también podrían ser t's de Student u otras) como necesitemos. Estas variables son independientes entre sí y las llamamos $(Y_1, \ldots, Y_n)$
2. En segundo lugar, queremos conseguir que estas variables estén correladas según una matriz de correlaciones dada. Para hacer esto observamos que si Ω es una matriz de correlaciones, admite una descomposición de la forma

$$\Omega = \text{Chol} \times \text{Chol}^{\text{traspuesta}}$$

donde Chol es una matriz triangular inferior. En ese caso las variables $(Z_1, \ldots, Z_n)$ definidas mediante

$$(Z_1, \ldots, Z_n) = (Y_1, \ldots, Y_n) \times \text{Chol}^{\text{traspuesta}}$$

están correladas según Ω.

A partir de estas normales ya correladas podemos obtener una tabla de posibles rendimientos tanto de los activos como de los derivados, y a partir de ahí obtenemos la tabla de rendimientos de la cartera sobre la cual podemos calcular los estadísticos de riesgo que deseemos.

EJERCICIO 6.3. *Implementar lo anterior en Excel, calculando el riesgo asociado a una cartera que consiste de 20 acciones de X, 25 acciones de Y y 10 opciones sobre X donde la correlación entre X e Y es* $\rho = 40\,\%$, $\Delta = 0,09$, $\Gamma = 0,006$, *usando la aproximación cuadrática. Compararlo con los resultados obtenidos al hacer la aproximación lineal.*

EJERCICIO 6.4. *En una cartera formada sólo por acciones, con una matriz de correlación dada, calcular el VaR por la metodología descrita inicialmente y a continuación comparar con el resultado obtenido mediante simulación Montecarlo.*

6.6. Otros modelos. Una de las grandes ventajas del modelo normal es que en ocasiones nos permite obtener fórmulas cerradas para nuestros estadísticos. Una vez que renunciamos a obtener fórmulas cerradas y que aceptamos que tenemos que usar Montecarlo, pasamos a sentirnos más libres para no usar el modelo normal y adoptar otros modelos.

Es bastante frecuente utilizar t's de Student en lugar de normales para modelizar los saltos. Esto es así porque la distribución t de Student es bastante similar a una normal, pero su principal diferencia es que las colas (los eventos extremos) de la t son más pesadas (tienen más probabilidad). Aparentemente, en muchas ocasiones los eventos extremos tienen más peso que el que les asigna la normal.

Veamos cómo simular con t's de Student. Todo es análogo al caso de la simulación mediante normales, tan sólo hay que tener cuidado en cómo correlarlas. Veamoslo con cierto detalle.

Una t de Student con n grados de libertad se define como

$$t^{(n)} = \frac{Z}{\sqrt{\frac{Z_1^2+Z_2^2+\cdots Z_n^2}{n}}}$$

donde $Z, Z_1, \ldots, Z_n$ son normales de media 0 y varianza 1 independientes.

Así definida, $t^{(n)}$ tiene media 0 y desviación típica $\sqrt{\frac{n}{n-2}}$. Es habitual considerar su normalización

$$t^{\hat{(n)}} = \frac{t^{(n)}}{\sqrt{\frac{n}{n-2}}}$$

que sigue teniendo media 0 y ahora tiene desviación típica 1.

Cuando n tiende a infinito, $t^{(n)}$ tiende a la normal.

Para generar una muestra de m $t^{\hat{(n)}}$'s de Student normalizadas correladas según una matriz de correlaciones dada procedemos como sigue.

1. Generamos primeramente m normales estándar $Z_1, \ldots, Z_m$ correladas según la matriz de correlaciones deseada, usando la descomposición de Cholesky de la matriz como vimos anteriormente.
2. Obtenemos una muestra Y de una χ-cuadrado con n grados de libertad (recordemos que χ-cuadrado se define como $Z_1^2 + \cdots + Z_n^2$) independiente de las normales anteriores.
3. Para todo $1 \leq j \leq m$ definimos ahora

$$t_j^{(n)} = \frac{Z_j}{\sqrt{\frac{Y}{n}}}.$$

   Las $t_j^{(n)}$ así definidas están correladas según la matriz deseada.

4. Finalmente, dividimos cada una de ellas por $\sqrt{\frac{n}{n-2}}$ para obtener las $\hat{t_j^{(n)}}$'s con la misma correlación.

Existe otra forma de hacer esto, que nos da una muestra correlada con "casi" la matriz deseada. Esta técnica se usa para correlar funciones de distribución en general, no sólo t's de Student. Se conoce como correlación mediante cópula normal.

1. Generamos una muestra de normales independientes y a partir de ellas generamos una muestra de normales correladas con la matriz de correlación deseada.
2. Pasamos de las muestras de normales a muestras de uniformes
3. Pasamos de las muestras de uniformes a las muestras de t's de Student (o a cualquier otra distribución deseada).

Las muestras finales no están correladas exactamente según la matriz deseada, aunque habitualmente la aproximan bastante.

Para pasar de la distribución normal a la uniforme, o de la uniforme a la t de Student, nótese que si X es una variable aleatoria con función de distribución F entonces $U = F(X)$ es una variable aleatoria uniforme, y que recíprocamente si U es una variable aleatoria uniforme entonces $X = F^{-1}(U)$ es una variable aleatoria con función de distribución F.

7. Estres de Correlaciones

En muchas ocasiones resulta interesante forzar el modelo con escenarios en los que las correlaciones no son las capturadas realmente de mercado, sino otras que los expertos consideran posibles en condiciones inusuales pero no imposibles.

En ese caso los expertos típicamente son capaces de captar las correlaciones 2 a 2, pero no las que involucran mayor número de activos. El resultado es que la "matriz de correlaciones" que sugieren no es tal, normalmente porque no es definida positiva.

Veamos cómo podemos resolver este problema.

Ya hemos mencionado anteriormente que si una matriz de correlaciones entre n activos Ω admite una descomposición de la forma

$$\Omega = \mathrm{Chol} \times \mathrm{Chol}^t$$

y si $(Y_1, \ldots, Y_n)$ son normales estándar independientes, entonces

$$(Z_1, \ldots, Z_n) = (Y_1, \ldots, Y_n) \times \mathrm{Chol}^t$$

son normales estándar correladas según la matriz Ω.

Este resultado sigue siendo cierto para cualquier descomposición

$$\Omega = AA^t.$$

Es decir, también en ese caso se tiene que

$$(Z'_1, \ldots, Z'_n) = (Y_1, \ldots, Y_n) \times A^t$$

son normales estándar correladas según la matriz Ω.

Supongamos en particular que Ω fuera diagonalizable. En ese caso

$$\Omega = UDU^t$$

donde U es una matriz unitaria real $(UU^t) = I$ y D es la matriz diagonal de los autovalores. Si Ω es una matriz de correlaciones, es definida positiva, de manera que sus autovalores son todos positivos, por lo que tiene sentido hablar de la matriz $\sqrt{D}$. De esta forma podemos escribir

$$\Omega = U\sqrt{D}\sqrt{D}U^t = \left(U\sqrt{D}\right)\left(\sqrt{D}U\right)^t$$

y por tanto

$$(Z'_1, \ldots, Z'_n) = (Y_1, \ldots, Y_n) \times \left(\sqrt{D}U\right)^t$$

serían normales estándar correladas según la matriz Ω.

Volvamos ahora a nuestro problema. Nos dan una matriz Ω "de correlaciones" que intenta capturar las correlaciones 2 a 2, pero que típicamente no lo hace 3 a 3, 4 a 4, etc y por tanto habitualmente no es definida positiva. Buscamos encontrar una auténtica matriz de correlaciones lo más cercana posible a Ω.

Antes de hacer esto, recordemos en primer lugar que una matriz Ω es de correlaciones si es simétrica, definida positiva y con unos en la diagonal.

Por ser simétrica, Ω diagonalizable; es decir, se puede escribir como

$$\Omega = UDU^t$$

donde

$$D = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}$$

es la matriz diagonal de autovalores y U es una matriz unitaria en la que las columnas son una base ortonormal de autovectores de Ω.

Por ser definida positiva, los autovalores son todos positivos.

Puesto que tiene unos en la diagonal, si Ω admite una descomposición de la forma

$$\Omega = AA^t$$

las filas de A deben tener norma cuadrática igual a 1.

Buscamos pues una matriz Ω' de correlaciones lo más cercana posible a Ω. Podemos proceder de la siguiente manera.

Puesto que Ω es simétrica, será diagonalizable, y la escribimos como

$$\Omega = UDU^t$$

donde

$$D = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}$$

es la matriz diagonal de autovalores, ordenados decrecientemente ($\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$). Puesto que Ω no era definida positiva, en general tendremos que algunos de los autovalores serán negativos. Si Ω estaba cercana a una matriz de correlaciones, los autovalores negativos serán pequeños en valor absoluto.

Entonces, para buscar una matriz que sea realmente de correlaciones en primer lugar sustituimos todos los autovalores negativos por 0's y pasamos a una matriz

$$D' = \begin{pmatrix} \lambda_1 & & & & & \\ & \ddots & & & & \\ & & \lambda_k & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}$$

Puesto que todos los autovalores de D' son positivos, tiene sentido definir

$$\sqrt{D'} = \begin{pmatrix} \sqrt{\lambda_1} & & & & & \\ & \ddots & & & & \\ & & \sqrt{\lambda_k} & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}$$

Recordemos que buscamos una matriz A tal que $\Omega' = AA^t$ sea una matriz de correlaciones y donde Ω' sea lo más próxima posible a Ω .

Definimos entonces

$$B = U\sqrt{D'}$$

Así definida, BB^t es una matriz simétrica definida positiva, pero en general no tendrá 1's en la diagonal. Para lograr esto debemos normalizar las filas de B, de manera que pasamos a la matriz A definida que se construye sin más que dividir cada fila de B por la norma 2 de dicha fila.

Entonces la matriz
$$\Omega' = AA^t$$
es simétrica, (semi)definida positiva, y con unos en la diagonal. Además es la más cercana a Ω con esas características, por lo que es la matriz que usaremos para correlar.

EJERCICIO 7.1. *Calcular el VaR de una cartera que tiene una "matriz de correlaciones" dada.*

8. Bonos: TIR

Se define la TIR (Tasa Interna de Rendimiento) como el factor de descuento (constante a lo largo de la vida del bono) que daría un precio para el bono igual a su precio actual. Veamos un ejemplo con el bono anterior.

Buscamos cuál sería el factor de descuento r constante que haría

$$2e^{-r\cdot 0,5} + 2e^{-r\cdot 1} + 2e^{-r\cdot 1,5} + 105e^{-r\cdot 2} = 101,89$$

Esto nos da un valor de $4,4\,\%$ para la TIR.

En general, si tenemos un bono que paga cantidades c_i en los instantes t_i y que vale B a día de hoy, la TIR del bono y verifica

$$B = \sum_{i=1}^{n} c_i e^{-yt_i}.$$

Observad que el rendimiento de la TIR se supone constante en el tiempo, es decir se correspondería a una curva cupón cero plana. La TIR de un bono es el nivel al que tendría que estar una curva cupón cero plana que nos valorara bien dicho bono.

9. Bonos: Duración

La *duración* de un bono se puede ver como una media, adecuadamente ponderada, de cuánto tiempo ha de transcurrir hasta recibir un pago. Su interés viene de que la duración mide (aproxima) la sensibilidad de un bono a la TIR, y por tanto a movimientos planos en la curva cupón cero.

En concreto, dado un bono que paga cantidades c_i en los instantes t_i, que vale B a día de hoy y cuya TIR es y, se define su duración como

$$D = \frac{\sum_{i=1}^{n} t_i c_i e^{-y t_i}}{B} = \sum_{i=1}^{n} t_i \frac{c_i e^{-y t_i}}{B}.$$

Puesto que $\sum_{i=1}^{n} \frac{c_i e^{-y t_i}}{B} = 1$, vemos que podemos interpretar la duración como la media de los instantes en los que vamos a recibir pagos, ponderado cada uno de estos instantes por el valor que tendría hoy el pago en ese instante si supusiéramos una curva de tipos plana.

El interés de la duración proviene del siguiente hecho (inmediato de comprobar)

$$\frac{dB}{dy} = -BD,$$

o en otra forma

$$\frac{\frac{dB}{dy}}{B} = -D$$

Utilizando este hecho y el desarrollo de Taylor de orden 1, podemos aproximar

$$\Delta B \sim -BD\Delta y$$

10. Bonos: Convexidad

Si deseamos una mejor aproximación a ΔB podemos utilizar el desarrollo de Taylor de orden 2. Para ello, supuesto un bono con las mismas características de antes, definimos su *convexidad* como

$$C = \frac{1}{B}\frac{d^2 B}{dy^2} = \frac{\sum_{i=1}^{n} c_i t_i^2 e^{-y t_i}}{B}$$

De esa manera, de nuevo es inmediato comprobar que

$$\frac{\Delta B}{B} \sim -D\Delta y + \frac{1}{2}C(\Delta y)^2.$$

Capítulo 7

Riesgo de Crédito

Otra de los riesgos importantes en los mercados es el llamado *riesgo de crédito*. Es el riesgo asociado a que tus deudores en una transacción cometan *impago* (default).

1. Clasificaciones crediticias

Las grandes compañías, así como los estados, tienen una *clasificación crediticia* (credit rating) concedida por alguna de las agencias que se dedican a ello (las principales son Moody's y S&P). Las clasificaciones de Moody's son Aaa, Aa, A, Baa, Ba, B, Caa y las correspondientes de S& P son AAA, AA, A, BBB, BB, B, CCC. Estas categorías a su vez se subdividen para mayor precisión.

1.1. Riesgos históricos de impago. Uno de los datos necesarios para entender el riesgo de crédito es una tabla histórica de riesgos de impago. Por ejemplo, la tabla siguiente nos da, en porcentaje, las probabilidades (acumuladas) históricas (en el periodo 1970-2003) de que las compañías con determinada clasificación crediticia hagan default al cabo de un determinado número de años.

Tiempo en Años	1	2	3	4	5	7	10	15	20
Aaa	0	0	0	0,04	0,12	0,29	0,62	1,21	1,55
Aa	0,02	0,03	0,06	0,15	0,24	0,43	0,68	1,51	2,70
A	0,02	0,09	0,23	0,38	0,54	0,91	1,59	2,94	5,24
Baa	0,2	0,57	1,03	1,62	2,16	3,24	5,10	9,12	12,59
Ba	1,26	3,48	6,00	8,59	11,17	15,44	21,01	30,88	38,56
B	6,21	13,76	20,65	26,66	31,99	40,79	50,02	59,21	60,73
Caa	23,65	37,20	48,02	55,56	60,83	69,36	77,91	80,23	80,23

Obsérvese que son las probabilidades *acumuladas*. A partir de ellas es inmediato obtener las respectivas probabilidades en cada periodo.

1.2. Tasas de recuperación. Si una compañía quiebra, y por tanto no puede hacer frente a sus obligaciones, sus acreedores pueden recuperar parte de la deuda con cargo a los activos de la compañía. En el proceso de recuperación de la deuda, los acreedores se dividen en diferentes categorías, y cada una de las categorías recibe un porcentaje diferente de su deuda.

2. Riesgo de impago a partir de los precios de los bonos

Veamos cómo se calcula la probabilidad (en el mundo *neutral al riesgo*) de impago de una compañía en función del precio de sus bonos.

Recordemos que muchas de las grandes compañías emiten bonos para financiarse. Los bonos emitidos por ciertos estados (entre ellos España) tienen la máxima calidad crediticia. Se supone que nunca van a hacer impago y el tipo de interés que pagan se conoce como *interés sin riesgo*. Puesto que el mercado considera que una compañía sí tiene riesgo de impago, la compañía tendrá que ofrecer un mayor tipo de interés (es decir, tendrá que venderlo más barato) por un bono de las mismas características que las del estado si quiere venderlo, ya que los inversores necesitan alguna compensación por el riesgo adicional que asumen.

Viendo por tanto los precios de los bonos y suponiendo que se puede estimar razonablemente bien la tasa de recuperación, se pueden deducir las probabilidades (de nuevo en el mundo *neutral al riesgo*) de impago de las distintas compañías.

3. Correlaciones entre los impagos

A la hora de valorar derivados de crédito, o a la hora de examinar el riesgo de crédito de una cartera compuesta por bonos de distintas compañías, es fundamental estudiar la correlación entre las probabilidades de impago de las distintas compañías involucradas.

Se observa que las probabilidades de quiebra de distintas compañías no son independientes. Esto resulta natural: por ejemplo, si las condiciones económicas empeoran, el riesgo de impago de todas las compañías aumenta. O, si hay dos empresas con mucha dependencia mutua, la quiebra de una de ellas aumenta notablemente la probabilidad de quiebra de la segunda. Compañías del mismo sector, o implantadas en la misma área geográfica, también pueden tener riesgos de quiebra fuertemente correlados entre sí.

Naturalmente esto afecta sensiblemente al riesgo de una cartera formada por bonos de distintas compañías. Es fácil darse cuenta de que dicho riesgo aumenta cuanto mayor sea la correlación entre las probabilidades de quiebra de las compañías.

El estudio de las correlaciones entre las distintas posibilidades de quiebra no es elemental y probablemente no esté aún bien entendido.

Normalmente se sigue un modelo de cópula normal para las correlaciones.

4. Derivados de crédito

Hasta hace relativamente poco, no había en el mercado instrumentos útiles para cubrirse del riesgo de crédito. Desde finales de los 90, el mercado de derivados de crédito ha ido creciendo espectacularmente tanto en sofisticación de los derivados como en el volumen de transacciones realizadas con ellos.

4.1. Credit Default Swaps. El derivado más habitual para protegerse del riesgo de crédito son los Credit Default Swap (CDS). Esencialmente un CDS es un seguro en caso de quiebra. La situación habitual es que una compañía A ha emitido bonos, y la compañía B ha comprado uno de estos bonos por un principal de digamos 10 millones de euros. Para cubrirse del riesgo de que A quiebre y no cumpla su compromiso al vencimiento del bono, B compra de una tercera compañía C un CDS sobre A, por el total de los 10 millones de euros con un coste de digamos 100 puntos básicos (un 1 %) anuales.

Entonces, si A no quiebra, B paga 100.000 euros anuales a C por el CDS. Si A quebrara, B tiene el derecho a venderle el bono de A a C por el principal (es decir, por los 10 millones de euros).

El efecto es que C está funcionando como una aseguradora que asegura a B del riesgo de que A quiebre.

5. CDO's

Un CDO (*collateralized debt obligation* es una manera habitual de transferir el riesgo en el que a partir de una cartera formada por varios instrumentos de deuda se crean instrumentos con perfiles de riesgo muy heterogéneos. La situación habitual es la siguiente: Una compañía A tiene una cartera formada por distintos bonos, cada una con distinto probabilidad de impago (o la compañía es El Corte Inglés y los bonos son los distintos pagos realizados por tarjeta de crédito por sus clientes y que aún no se han hecho efectivos, o la compañía es un banco y los bonos son los distintos créditos pendientes de pago concedidos a sus clientes). Si la compañía no desea conservar esa cartera por el riesgo que lleva asociado, una posibilidad es crear un CDO. Este consiste en distintas "tranches", cada una de las cuales asume sucesivos porcentajes del riesgo. Es frecuente que estos porcentajes salten de 3 % en 3 %. Supongamos esa situación. En ese caso la tranche 1 es un instrumento emitido por la compañía con las siguientes condiciones: si ninguno de los bonos comete impago, la tranche 1 pagará un cierto interés (bastante alto en el caso de la tranche 1). Si un 1 % de los bonos comete impago, el principal (y los intereses) de la tranche 1 se vera reducido en un tercio. Si el 3 % de los bonos comete impago, el principal de la tranche 1 se reduce a 0. La tranche 2 funciona igual, pero ahora asume los riesgos del 3 % al 6 %. Es decir, si menos de un 3 % de los bonos comenten impago, la tranche 2 paga todo lo prometido. Si el 4,5 % de los bonos cometen impago, la tranche 2 pierde la mitad de su valor, y si el porcentaje de los bonos que comenten impago es superior al 6 % la tranche 2 pasa a valer 0. Y así sucesivamente. Normalmente a partir de cierto porcentaje (digamos el 25 %) lo que resta ya tiene máxima clasificación crediticia (AAA) y no se subdivide más sino que se vende junto, o lo retiene la compañía A puesto que se considera libre de riesgo. El resultado es que las tranches iniciales, sobre todo la primera, pasan a ser instrumentos de muy alto riesgo, y el riesgo de las sucesivas tranches va disminuyendo.

Valorar las distintas tranches de un CDO es un problema no trivial, sobre todo en el caso de que la cartera está formada por un alto número de bonos/deudores. A la hora de valorar, la correlación entre las distintas probabilidades de impago es fundamental. Se suelen correlar con el modelo de cópula normal, aunque se sabe que este modelo no captura el sentir del mercado.

Bibliografía

[1] *Quantitative Modelling of Derivative Securities*, M. Avellaneda and P. Laurence. Chapman and Hall 2000.

[2] F. Black and M. Scholes, The valuation of option contracts and a test of market efficiency, *J. Finance* **27** (1972), 399–417.

[3] F. Black and M. Scholes, The pricing of options and corporate liability, *J. Political Economy*, **81** (1973), 637–659.

[4] R. Dana and M. Jeanblanc, *Financial Markets in Continuous Time*, Springer 2003.

[5] S. Dineen, *Probability Theory in Finance: A Mathematical Guide to the Black-Scholes Formula*, American Mathematical Society, Graduate Series in Mathematics, 2005.

[6] D. Duffie, *Security Markets: Stochastic Models*, Academic Press 1988.

[7] R. J. Elliot and P. E. Kopp, *Mathematics of Financial Markets*, Springer, 1999.

[8] J. C. Hull, *Options, Futures and Other Derivatives*, 5ï£¡ Ed. Prentice Hall 2003.

[9] I. Karatzas and S. E. Shreve, *Brownian motion and stocastic calculus*, Springer 1991.

[10] P. E. Kopp, *Martingales and stochastic integrals*, Cambridge University Press 1984.

[11] R. Merton, The theory of rational option pricing, *Bell. J. Econ. Manage. Sci.,* **4** (1973), 141–183.

[12] J. Neveu, *Discrete parameter martingales*, North-Holland 1975.

[13] W. Shachermayer, A Hilbert space proof of the fundamental theorem of asset pricing in finite discrete time, *Insurance Math. Econ.* **11** (1992) 249–257.