Riesgo e Incertidumbre

(Cap. 9 del Tratado de Microeconomía)

Definiciones

El riesgo es un concepto generalmente mal comprendido y asociado a diferentes nociones, muchas veces contradictorias, incluso en economía. El primer autor en diferenciar estos dos conceptos en forma clara fue el economista estadounidense de la Universidad de Chicago **Frank Knight**, quien, hace unos 90 años, en su libro titulado *Risk, Uncertainty, and Profit* (Riesgo, incertidumbre y beneficio) definió a la *incertidumbre como un riesgo no medible y no computable*. Más recientemente, en 1985, **Oliver Williamson** otorga a la incertidumbre un rol preponderante como uno de los tres atributos de la transacción que las firmas deben considerar para establecer la mejor estructura de gobernanza posible (mercado, integración vertical o contratos), para intentar reducir costos de transacción.

El término **gobernanza** viene utilizándose desde la década de los 1990s para designar la eficacia, calidad y buena orientación de la intervención del Estado, que proporciona a éste buena parte de su legitimidad en lo que a veces se define como una "nueva forma de gobernar" en la globalización del mundo posterior a la caída del muro de Berlín (1989).

Teniendo estas ideas en mente, podemos definir y separar los conceptos de riesgo e incertidumbre según dos aspectos clave: 1) la **posibilidad de medirlos**, **diversificarlos** y **mitigarlos**, y 2) la **capacidad de los agentes de actuar** sobre ellos. Ejemplo sencillo: ignoramos cuál será el resultado de tirar un dado, pero sabemos que será uno de los resultados 1-2-3-4-5 o 6. *Esto no es incertidumbre*, sino *riesgo*.

El riesgo es un elemento inherente a cualquier fenómeno. A su vez, cuando los factores son múltiples y se interrelacionan, generan incertidumbre. En matemática de **teoría del caos**, se conoce como “comportamiento estocástico de sistemas deterministas”, que es una manera complicada de explicar algo sencillo: es relativamente fácil explicar fenómenos conocidos aislados, pero en la medida en que se trata de fenómenos que ocurren de forma simultánea y que inciden unos sobre otros, es sumamente difícil pronosticar el resultado. Un ejemplo es el clima: explicar cómo se produce la lluvia es sencillo, pero **pronosticar el clima en periodos largos es complejo** porque cada fenómeno interactúa con otros que a su vez lo afectan. Así, los monzones en India están relacionados con la presencia del fenómeno de El Niño en la costa oeste de América del Sur. En lo financiero ocurre algo similar. Podemos suponer que el comportamiento de las tasas o el riesgo de una inversión dependen de ciertos factores concretos, pero al final, temas tan aparentemente lejanos como un huracán que afecta las cosechas de tomate, inciden en la inflación que, a su vez, mueve la tasa de interés, o un plan de rescate en una economía pequeña en Europa puede afectar la percepción de riesgo financiero mundial. *La incertidumbre es inherente al mundo de las finanzas*.

El riesgo es **no sistémico y diversificable y, por ende, puede ser reducido**. Es no sistémico porque un agente económico puede tener injerencia sobre él; diversificable porque puede ser medido, diversificado y mitigado. **Esto es lo que hacen los productores agropecuarios al operar en varias zonas distintas, al contratar seguros o al fijar precios a futuro: reconocen un riesgo a nivel empresa e intentan mitigarlo con este tipo de actividades para reducir su posible impacto**. En definitiva, es imposible que el productor tome decisiones productivas sin asumir un cierto riesgo, lo cual no implica que no sea posible lograr reducir este riesgo por medio de distintas herramientas. La mejor combinación de estrategias dependerá de cada productor porque las posibilidades de enfrentar situaciones de riesgo y la preferencia por mayores o menores niveles de riesgo dependen de cada individuo.

En contraposición, **la incertidumbre es sistémica y no es diversificable** y, por ende, no puede ser **medida** ni **mitigada**. Surge por el solo hecho de llevar a cabo una actividad; está determinada por el sistema económico, es decir, que los agentes no pueden actuar sobre ella directamente y será mayor o menor dependiendo de varios aspectos. Tal vez el más importantes es el **ambiente institucional**, que Douglass North define como un conjunto de restricciones creadas por el hombre que estructuran las interacciones políticas, económicas y sociales, que pueden ser formales (constituciones, leyes o derechos de propiedad) o informales (sanciones, tabúes, costumbres o tradiciones), o sea, las reglas de juego que rigen en una sociedad. En **ambientes institucionales frágiles**, donde se dan permanentes cambios de las reglas de juego, con bajo **enforcement** (logro del cumplimiento) de la ley, con mayor dificultad para prever situaciones conflictivas que redundarán en una mayor incompletitud de los contratos y, por consiguiente, con grandes incentivos para actuar oportunistamente, la incertidumbre será mayor y, consecuentemente, los costos de transacción también serán mayores.

Esta distinción de conceptos resulta de gran importancia debido a que **las innovaciones más frecuentes (las tecnológicas y las organizacionales) se enfocan hacia el riesgo**, mientras que **la incertidumbre solo puede ser reducida mediante innovaciones en el ambiente institucional, sobre el cual los agentes individuales no tienen injerencia por sí mismos**, sino que son las sociedades en su conjunto las que deben operar para modificarlo.

En consecuencia, adoptamos las siguientes definiciones:

Riesgo – No sabemos qué sucederá pero al menos se conoce su distribución de probabilidad

Incertidumbre – No sabemos qué sucederá ni tampoco sabemos cuál sería su posible distribución de probabilidad

Knight: *La incertidumbre debe ser tomada en un sentido radicalmente distinto de la noción familiar de riesgo, de la que nunca se ha separado adecuadamente.... El hecho esencial es que «riesgo» significa, en algunos casos, una cantidad susceptible de medición, mientras que en otros casos no es así; y hay diferencias cruciales en el funcionamiento de los fenómenos dependiendo de cuál de ambos está realmente presente y operativo....* ***Se verá que una incertidumbre medible, o "riesgo", como vamos a utilizar el término, es tan diferente de un riesgo inconmensurable, que realmente es mejor pensar que no se trata de incertidumbre****.*

Luego, en términos de Knight, La incertidumbre no es medible ni tampoco es posible calcularla, mientras que el riesgo sí lo es.

<u>Variables aleatorias y distribuciones de probabilidad</u>

Estos temas son un repaso de cuestiones que son analizadas en Estadística y Probabilidad.

En probabilidad y estadística, una **variable aleatoria (v.a.)**, o **variable estocástica** es una variable cuyo valor está sujeto a variaciones debido a la probabilidad (es decir, al azar, en sentido matemático).

Dada una variable aleatoria no es posible conocer con certeza el valor que tomará ésta al ser medida o determinada, aunque sí se conoce que **existe una distribución de probabilidad asociada al conjunto de valores posibles**. Por ejemplo, en una epidemia de cólera, se sabe que una persona cualquiera puede enfermar o no (suceso), pero no se sabe cuál de los dos sucesos va a ocurrir. Solamente se puede decir que existe una probabilidad de que la persona enferme.

Función de distribución La distribución de probabilidad de una variable aleatoria es una función que asigna a cada suceso definido sobre la variable aleatoria, la probabilidad de que dicho suceso ocurra. La distribución de probabilidad está definida sobre el conjunto de todos los sucesos, cada uno de los sucesos es el rango de valores de la variable aleatoria.

Supongamos que se lanzan dos monedas al aire. El **espacio muestral**, esto es, el conjunto de resultados elementales posibles asociado al experimento, es

$\Omega = \{cc, cx, xc, xx\}$ (donde c representa "sale cara" y x, "sale cruz").

Podemos asignar entonces a cada suceso elemental del experimento el número de caras obtenidas. De este modo se definiría la **variable aleatoria** *x* como la función

$$X: \Omega \rightarrow R$$

dada por: $cc \rightarrow 2;$
$cx, xc \rightarrow 1;$
$xx \rightarrow 0.$

Otro ejemplo. Si arrojamos una moneda, podemos definir una v.a. X tal que X=47 si sale cara y X=35 si sale cruz. Asociada a cada v.a. X existe una **función de distribución acumulada** F tal que:

[1] $$\boldsymbol{F(x) = Pr[X \leq x]}.$$

Continuando con el último ejemplo, si la moneda no está sesgada, la distribución de X podría venir dada por:

[2] $$F(x) = \begin{cases} 1 \text{ para } x \geq 47 \\ 0.5 \text{ para } 47 > x \geq 35 \\ 0 \text{ para } x < 35 \end{cases}$$

Una función de distribución debe tener las propiedades siguientes, como es el caso de [2]:

$F(\infty)=1;\ F(-\infty)=0;\ F(x)$ debe ser monótona no decreciente en x.

Las v.a. pueden ser discretas o continuas. Si la v.a. X es discreta, sólo puede tomar un número finito o infinito numerable de valores.

La **función de densidad de probabilidad** f asociada con X es

[3] $f(x) = Pr[X=x]$

Naturalmente, las probabilidades deben ser no negativas y sumar uno. Luego, tenemos que $\mathbf{1 \geq f(x) \geq 0}$ y $\mathbf{\sum_x f(x) = 1}$.

Podemos hallar la probabilidad de que la v.a. esté comprendida en un pequeño intervalo, Pr[x≤X≤x+h]. Dividiendo esta probabilidad por la longitud del intervalo y tomando el límite cuando su longitud tiende a cero, se obtiene la **función de densidad de probabilidad**:

[4] $$f(x) = \lim_{h\to 0} \frac{Pr[x\leq X\leq x+h]}{h} = \lim_{h\to 0} \frac{F(x+h)-F(x)}{h} = F'(x)$$

La función de densidad de probabilidad también debe satisfacer dos condiciones:

$f(x)\geq 0$

$\int_{-\infty}^{\infty} f(x)\, dx = 1.$

Media y varianza Una v.a. puede ser completamente caracterizada por su función de distribución. Empero, es útil resumir la tendencia central o promedio de la v.a. mediante un número real. La medida más importante de tendencia central es la media, también denominada **valor esperado de una v.a.** x y denotada mediante $E[x]$, definida como

[5] $E[x]= \sum_x x f(x)$ si x es discreta

$E[x]=\int_{-\infty}^{\infty} x f(x)\, dx$ si x es continua.

La media o esperanza matemática es un operador **lineal**: Si X e Y son dos variables aleatorias, y a, b son números: $E[aX + bY] = a E[X] + b E[Y]$.

Ejemplos a) Si una persona compra un número en una rifa, en la que puede ganar $5.000 ó un segundo premio de $2000 con probabilidades de: 0.001 y 0.003. ¿Cuál sería el precio justo a pagar por el boleto? Respuesta: E(x) = 5000 · 0.001 + 2000 · 0.003 = $11

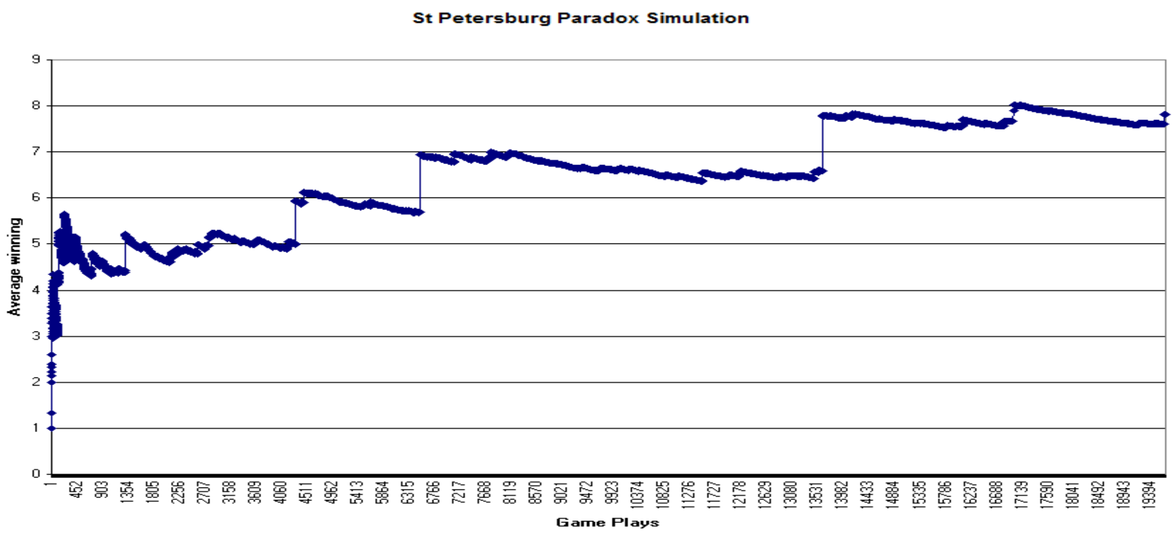
b) Tomemos la siguiente apuesta. Una moneda sin sesgo es arrojada hasta que aparezca "cara". Ustedes ganan $1 si aparece en la primera tirada, $2 si aparece en la segunda, $4 si lo hace en la tercera, y en general, $\$2^{n-1}$ si aparece en la n-ésima. Denotando lo ganado como una v.a. x, la probabilidad de ganar $\$2^{n-1}$ es $(½)^n$. Luego, el valor esperado de x es una suma divergente:

$E[x]=\sum_n 2^{n-1}(½)^n = ½ + ½ + ½+\ldots \to\infty$

Ésta es la conocida **Paradoja de San Petersburgo**. La Paradoja desafía la idea de que la gente evalúa a los proyectos aleatorios de acuerdo con su valor esperado. La Paradoja planteaba la situación en los términos siguientes: una moneda insesgada sería arrojada hasta que aparezca una cara; si la primera cara aparece en la n-ésima tirada, el pago será igual a 2^n ducados. ¿Cuánto uno estaría dispuesto a pagar por jugar este juego? La Paradoja, como hemos visto, es que el retorno esperado es infinito, pero aunque el pago esperado sea infinito

¡nadie supondrá, al menos intuitivamente, que la gente esté dispuesta a pagar una cantidad infinita de dinero para jugarlo!

Un gráfico típico de las ganancias promedio recibidas en una corrida de la lotería de San Petersburgo muestra cuán ocasionalmente los pagos recibidos conducen a un aumento de las ganancias promedio, en general muy reducido. Después de 20.000 jugadas de esta simulación la ganancia promedio por lotería se ubicó algo por debajo de 8 dólares. El gráfico muestra la paradoja. Por un lado, la pendiente positiva muestra que las ganancias promedio tienden hacia infinito, pero la lentitud del crecimiento (lentitud que se hace más patente con el progreso del juego) indica que será necesario un número tremendamente elevado de jugadas para lograr una ganancia modesta. Un matemático del siglo XVIII planteó una solución involucrando dos ideas que, desde entonces, significaron una profunda revolución en economía: primero, que la utilidad que la gente deriva de la riqueza, *U(W),* no depende linealmente de la riqueza (*W*) sino que se incrementa a una tasa decreciente – la famosa idea de la *utilidad marginal decreciente*, $U'(W)>0$, y $U''(W)<0$; que la evaluación de una persona de un proyecto riesgoso no consiste del rendimiento esperado del proyecto, sino más bien de la **utilidad esperada de ese proyecto**.



En el caso de San Petersburgo, el valor del juego para un agente (suponiendo riqueza inicial igual a cero) es:

$$E(U) = \sum_{i=1\ldots\infty} U(2^{n-1})\,(1/2)^n = U(1).(1/2) + U(2).(1/4) + \ldots. < \infty.$$

que Nicolaus Bernoulli conjeturó finito a causa del principio de la utilidad marginal decreciente (el documento fue publicado por su hermano Daniel Bernoulli en los *St. Petersburg Academy Proceedings* 1738)[1]. En consecuencia, la gente sólo estaría dispuesta a pagar un monto finito de dinero para jugarlo, aunque su retorno esperado sea infinito. Cabe mencionar que Bernoulli usó una función logarítmica $u(x)=ln\ (x)$. En 1728, otro matemático suizo, Gabriel Cramer, había llegado a una conclusión similar utilizando como función de utilidad la raíz cuadrada de x. Sin embargo, a diferencia de Bernoulli, utilizó como argumento la ganancia monetaria de la lotería en lugar de la riqueza total. Pero las soluciones de Cramer y

[1] D. Bernoulli, Specimen Theoriae Novae de Mensura Sortis. Traducción: "Exposition of a new theory on the measurement of risk" by L. Sommer (1954). Econometrica 22(1).

Bernoulli no son completamente satisfactorias, ya que el juego puede ser fácilmente cambiado para que reaparezca la paradoja (Karl Menger puntualizó en 1934 que se puede encontrar una variante de la paradoja de San Petersburgo para cualquier función de utilidad no acotada). Empero, la función *ln(x)* se hace infinito a medida que su argumento tiende a infinito, pero de manera muy lenta.

Podemos extender la fórmula de valor esperado en forma natural a *funciones* de una variable aleatoria. Si $u(x)$ es función de una v.a. x, la función $u(x)$ se convierte en una v.a. y su valor esperado será $E[u(x)]=\sum_x u(x)f(x)$ si x es discreta, y $E[u(x)]=\int_{-\infty}^{\infty} u(x)f(x)\,dx$ si x es continua.

En general, a menos que u sea una función lineal de x, se tendrá que $\boldsymbol{E[u(x)]\neq u(E[x])}$. Pero si $u=a+bx$, a y b constantes, $\boldsymbol{E[a+bx]=a+bE[x]}$. La linealidad se aplica también a dos o más v.a. Para dos v.a. x e y, tendremos $\boldsymbol{E[x+y]=E[x]+E[y]}$, sean o no independientes. Aquí cabe apreciar que dos v.a. x e y son **independientes** si y sólo si

$$Pr[x\leq x^o, y\leq y^o]=Pr[x\leq x^o]Pr[y\leq y^o] \text{ para todo par } x^o,y^o .$$

Si x e y son **independientes**, se tiene que $E[xy]=E[x]E[y]$.

Varianza Además del valor esperado de una v.a. como medida de su tendencia central, es importante la varianza que indica el grado de variabilidad de la v.a. La varianza de una v.a. x, denotada como $var[x]$, se define mediante: $var[x]=E[(x-\mu)^2]$, donde $\mu=E[x]$. Esta última ecuación puede también ser expresada de la forma siguiente:

$$var[x]=E[x^2-2x\mu+\mu^2]=E[x^2]-2\mu E[x]+\mu^2=E[x^2]-2\mu^2+\mu^2=E[x^2]-\mu^2 .$$

Esta expresión suele mencionarse de la siguiente forma: el momento centrado de segundo orden (la varianza) es igual al momento absoluto de segundo orden menos la media al cuadrado.

Propiedades de la varianza:

1.- Para dos constantes cualesquiera a y b, $var[a+bx]=b^2 var[x]$

2.- Si $x_1,\ldots x_n$ son v.a. independientes, $var[x_1+\ldots+x_n]=var[x_1]+\ldots+var[x_n]$.

Covarianza La covarianza entre x_i y x_j, igual a $cov[x_i,x_j]=E[x_ix_j]-\mu_i\mu_j$. Luego, es válido que $var[x_1+x_2]=var[x_1]+var[x_2]+2\,cov[x_i,x_j]$.

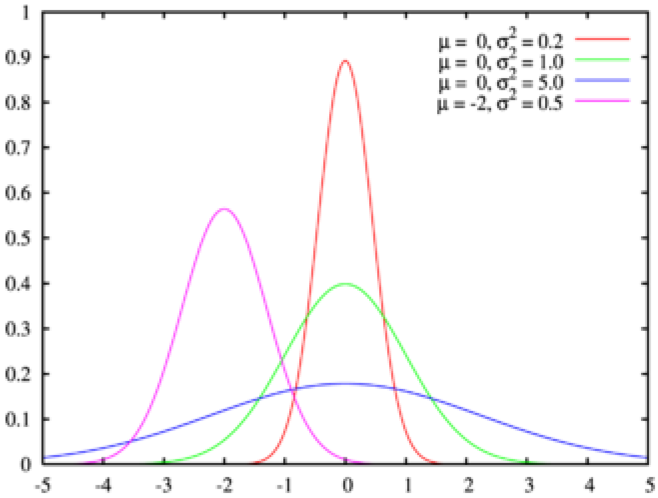
Generalmente, si $X=(X_1,\ldots,X_n)$ son variables aleatorias, y $\mu_i=E[X_i]$, entonces la matriz de covarianza Σ es la matriz cuya entrada (i,j) es la covarianza $\Sigma_{ij}=E[(X_i-\mu_i)(X_j-\mu_j)]$. La matriz de covarianza Σ es **simétrica**.

Ejemplo: diversificación Sea x el valor de una acción de una empresa. Si σ^2 es su varianza, una cartera consistente de n acciones de la empresa tendrá una varianza igual a $var[nx]=n^2\sigma^2$. Por otro lado, si ustedes invierten en acciones de n empresas distintas, $x_1,\ldots,x_n$ cuyos rendimientos son independientes. Si cada acción tuviera la misma varianza σ^2, la varianza de esta cartera diversificada sería igual a $var[x_1+x_2+\ldots+x_n]=n\sigma^2$. ***Han logrado reducir la varianza o riesgo de su cartera <u>diversificando</u>!***

La función de distribución normal

La importancia de esta distribución radica en que permite modelar numerosos fenómenos naturales, sociales y psicológicos. Mientras que los mecanismos que subyacen a gran parte de este tipo de fenómenos son desconocidos, por la enorme cantidad de variables incontrolables que en ellos intervienen, el uso del modelo normal puede justificarse asumiendo que cada observación se obtiene como la suma de unas pocas causas independientes.

La **distribución normal o de Gauss** es una familia importante de distribuciones continuas de probabilidad que es aplicada en diversos campos. Todo miembro de la familia puede ser definido con sólo dos parámetros de ubicación y escala: la media μ y la varianza σ^2, cuadrado del desvío estándar σ. **La distribución normal estándar** es la distribución normal con **media igual a cero y varianza igual a 1**. Carl Friedrich Gauss resultó asociado con este conjunto de distribuciones al analizar datos astronómicos mediante las mismas, y definió la ecuación de su función de densidad probabilística. A menudo es llamada la curva en campana a causa del gráfico de su función de densidad de probabilidad. Su importancia como modelo de los fenómenos cuantitativos en las ciencias naturales y de la conducta descansa en el **teorema central del límite** (*si S_n es la suma de n variables aleatorias independientes y de varianza no nula pero finita, entonces la función de distribución de S_n «se aproxima bien» a una distribución normal, para $n \to \infty$*). Varias mediciones, desde las psicológicas hasta las físicas y las sociales, pueden ser aproximadas por la distribución normal.



Función de densidad de probabilidad

La línea verde corresponde a la distribución normal estándar

La distribución normal también es importante por su relación con la estimación por mínimos cuadrados, uno de los métodos de estimación más simples y antiguos.

Algunos ejemplos de variables asociadas a fenómenos naturales que siguen el modelo de la normal son:

- Los caracteres morfológicos de individuos como la estatura;
- Los caracteres fisiológicos como el efecto de un fármaco;
- Los caracteres sociológicos como el consumo de cierto producto por un mismo grupo de individuos;
- Los caracteres psicológicos como el cociente intelectual;
- El nivel de ruido en telecomunicaciones;
- Los errores cometidos al medir ciertas magnitudes; etc.

La distribución normal también aparece en muchas áreas de la propia estadística. Por ejemplo, la distribución muestral de las medias muestrales es aproximadamente normal, cuando la distribución de la población de la cual se extrae la muestra no es normal (como consecuencia del *Teorema Central del Límite)*. La distribución normal es la más extendida en estadística y muchos tests estadísticos están basados en una "normalidad" más o menos justificada de la variable aleatoria bajo estudio. En probabilidad, la distribución normal aparece como el límite de varias distribuciones continuas y discretas.

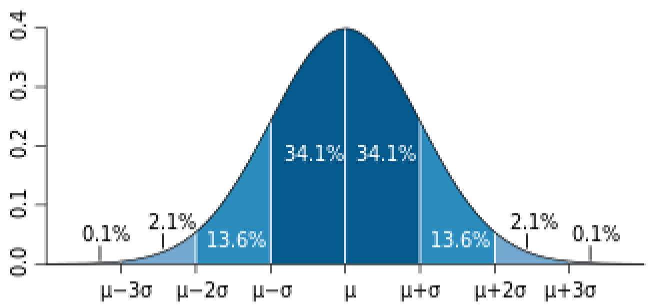
La función de densidad de probabilidad de la distribución normal es la ***función de Gauss***:

$$f(x) = ae^{-\frac{(x-b)^2}{2c^2}}$$

En esta expresión, todos los parámetros son positivos.

La función de densidad de probabilidad tiene ciertas propiedades notables como las siguientes:

1. Simetría en torno a su media $\boldsymbol{\mu}$
2. El modo y la mediana de la función son iguales ambos a la media μ.
3. Los puntos de inflexión de la curva se producen a la distancia de un desvío estándar de la media, e.d. en $\boldsymbol{\mu - \sigma}$ y en $\boldsymbol{\mu + \sigma}$.



Cerca de un 68% de los valores extraídos de una distribución normal se hallan dentro de la proximidad de un desvío estándar $\boldsymbol{\sigma}$ de la media $\boldsymbol{\mu}$. Alrededor de un 95% de los valores se encuentran dentro de dos desvíos estándar, mientras que un 99.7% lo hace dentro de tres desvíos estándar. *Esto conduce a una "regla empírica" conocida como "regla de 68-95-99.7".*

La aproximación de preferencias por el estado

El enfoque desarrollado en bolillas anteriores del programa puede ser generalizado en forma directa a fin de abarcar también la conducta bajo incertidumbre. Así como una manzana consumida hoy es algo diferente de una manzana consumida mañana, tomar un refresco en un día caluroso es diferente de hacerlo en un día frío. En problemas relacionados con la incertidumbre, es posible tratar al mismo bien físico consumido en distintos estados del mundo como bienes distintos. **Mediante este enfoque, la utilidad es definida en términos de mercancías estado-contingentes**. Un bien estado-contingente es un bien que sólo puede ser consumido si se verifica cierto estado del mundo. Por ejemplo, consideremos un contrato para entregar 100 l. de bebidas cola si la temperatura se ubica por encima de 25 ºC (pero nada en otros casos). Supongan que hay sólo dos estados posibles, y denotemos como W_1 y W_2 las cantidades de bienes estado 1 y estado 2- contingentes, respectivamente. W_1 y W_2 podrían ser vectores como si fueran canastas de mercancías, pero serán usados mayormente como escalares que representan a la riqueza o a un consumo compuesto. Si las probabilidades de ambos estados son respectivamente π_1 y π_2, con $\pi_1+ \pi_2=1$, las preferencias del consumidor pueden ser representadas por una función de utilidad como la siguiente:

$$U=U\,(W_1, W_2; \pi_1, \pi_2).$$

En tal caso la utilidad queda definida sobre el plan de consumo contingente (W_1, W_2). ¿Por qué están las probabilidades? Pues porque el valor de una mercancía estado-contingente depende de cuán probable es que ocurra ese estado. Si existen mercados completos en los que se puedan comprar bienes estado-contingentes a precios exógenos, el análisis de la elección de consumo bajo incertidumbre resulta formalmente equivalente al caso de certidumbre. La unidad consumidora elegirá W_1 y W_2 de modo de maximizar la utilidad sujeta a restricciones presupuestarias. Las funciones de demanda de bienes estado-contingentes resultantes satisfarán todos los teoremas ya derivados. **Pero el requerimiento de que existan mercados completos es muy exigente, ya que si hay *n* bienes diferentes y *s* estados posibles del mundo, tendría que haber *n.s* mercados separados**. Piensen ustedes en las grandes dimensiones del problema.

Kenneth Arrow demostró que el comercio de bienes estado-contingentes puede ser sustituido por el intercambio de **derechos estado-contingentes** (es decir, contratos financieros que dan lugar a distintas cantidades de dinero en estados distintos del mundo). Un conjunto completo de mercados sólo requiere **n mercados de bienes más s mercados de valores**. Luego, en lugar de ***n.s*** **mercados sólo se requieren** ***n+s*****.** Pero los mercados contingentes son difíciles de organizar (la especificación del estado y la medición son complejas). Sin embargo, existen algunos, como los **mercados de opciones.** Por ejemplo, el tenedor de una **opción put** solo ejercerá el Derecho de compra, si el precio de Mercado es inferior al precio de opción.

Pero esta aproximación puede resultar inabordable. Si no se comercian bienes estado-contingentes, hay que explorar alguna otra alternativa para tratar la existencia de incertidumbre.

Los consumidores como *timberos*

timba

1. f. Partida de juego de azar:
me voy a una timba que ha organizado un amigo.
2. Casa de juego, garito.

La teoría moderna de la probabilidad condujo a los economistas a hacer la hipótesis de que aún las creencias individuales pueden ser representadas mediante probabilidades personales o subjetivas, que adoptan la forma de una medida de probabilidad subjetiva aditiva $\mu(\ldots)$ en un espacio de estados S. En tal caso, **una canasta de pagos según el estado $(c_1,\ldots,c_n)$** será considerada como generando el resultado c_i con probabilidad $\mu(s_i)$, de modo que el individuo evaluará la canasta $(c_1,\ldots,c_n)$ de la misma forma en que evaluaría una apuesta en un casino que le depare los pagos $(c_1,\ldots,c_n)$ con probabilidades $(\mu(s_1), \ldots, \mu(s_n))$. **La hipótesis de que los individuos tienen tales creencias probabilísticas y que evalúan a las canastas de esta manera es denominada la hipótesis de sofisticación probabilística**, y permite una aplicación unificada de la teoría de la probabilidad al análisis de decisiones tanto bajo riesgo objetivo como bajo incertidumbre subjetiva.

Teoría de la Utilidad Esperada

La función de utilidad que vimos en la página 9, $U=U\ (W_1,\ W_2;\ \pi_1,\ \pi_2)$ es demasiado general para llegar a resultados docimables. Para imponerle más estructura hay algunos axiomas adicionales como los siguientes (von Neumann & Morgenstern)

1. Independencia del estado. Una perspectiva incierta que proporciona x en el estado 1 e y en el estado 2 es igualmente preferida a una perspectiva de y en el estado 1 y x en el estado 2, si la probabilidad de recibir x en ambos estados es la misma. La independencia del estado significa que las preferencias dependen de las probabilidades de los estados del mundo, no de los estados en sí mismos. *La validez de este supuesto depende del contexto; por ejemplo, en casos de seguro médico, las preferencias pueden depender del estado de salud aún si son cubiertos todos los gastos médicos.*

2. **Reducción de las loterías compuestas**. Si x es una perspectiva incierta consistente de *y* y *z* con probabilidades π y $1-\pi$, entonces la perspectiva que consiste de *x* y *z* con probabilidades π^* y $1-\pi^*$ es igualmente preferida que una perspectiva de *y* y *z* con probabilidades $\pi\,\pi^*$ y $(1-\pi)\pi^*$ e igualmente preferida a la perspectiva de *y* y *z* con probabilidades $\pi\,\pi^*$ y $(1-\pi)\,\pi^*$. Lo que el axioma afirma es que las preferencias de un consumidor por perspectivas inciertas dependen únicamente de las probabilidades de percibirlas y no de cómo se forman las probabilidades. *Si un consumidor tuviera una inclinación particular hacia el suspenso, este axioma sería violado.*

3. Continuidad. Si *x* es preferido a *y* y éste preferido a *z,* existe un cierto valor de probabilidad π tal que ***y*** será preferido a una **combinación incierta consistente de *x*** y ***z****, con x realizado con probabilidad* π *y z con probabilidad* $1-\pi$. Este axioma no sería válido si *x* es $2, *y* es $1, y *z* es la muerte. Por otra parte, a menudo la gente asume el riesgo de cruzar la calle de forma imprudente sólo para ganar unos segundos.

4. Independencia de alternativas irrelevantes. Si *x* es preferido a *y*, luego para cualquier *z* una perspectiva incierta consistente de *x* y *z* con probabilidades π y $1-\pi$ será preferida a la perspectiva incierta consistente de *y* y *z* con idénticas probabilidades. Si hay certidumbre, este axioma de independencia es un postulado fuerte, porque entre dos bienes pueden existir todo tipo de relaciones de complementariedad o de sustitución, ya se consuman en forma simultánea o sucesiva en el tiempo. Para perspectivas inciertas, empero, un individuo nunca recibirá ***x*** y ***z*** conjuntamente, o ***y*** y ***z*** conjuntamente. Luego, la presencia de z es improbable que afecte las preferencias por x e y. Este axioma no debe ser confundido con un axioma con el mismo nombre usado en la teoría de la elección social.

Teorema Bajo los axiomas [1]-[4], si los valores W_1 y W_2 con probabilidades respectivas π_1 y π_2 están dados, **es posible hallar una función de utilidad *u (…)*** – que será llamada la “función de utilidad elemental o de Bernoulli” – tal que

$$\boldsymbol{U\ (W_1,\ W2;\ \pi_1,\ \pi_2) = \pi_1 u\ (W_1) + \pi_2 u\ (W_2)}$$

Esta función *U (.)* es llamada **función de utilidad de Von Neumann-Morgenstern** en honor a la contribución pionera de estos autores en teoría de la decisión (1944). Como resultado se obtiene que las preferencias –expresadas por *U (.)* – **son el valor esperado de la función de utilidad elemental o de Bernoulli, aditivamente separable** en los resultados W_1 y W_2 y lineal en π_1 y π_2. La separabilidad es una consecuencia del axioma 4. Si no se cumple el axioma 4 de independencia del estado pero se mantienen los restantes, la utilidad puede ser expresada como $\pi_1 u_1\ (W_1) + \pi_2 u_2\ (W_2)$, donde u_1 y u_2 son funciones diferentes.

Utilidad cardinal y utilidad ordinal

Tal como en el caso bajo certidumbre, la función de utilidad no es nada más que una manera conveniente de representar preferencias. Si se tiene que **x≿y** todas las veces que *U(x) ≥U (y), U (.)* es una función de utilidad válida. Como *U(x) ≥U(y)* implica *F(U(x))≥F(U(y))* para cualquier transformación monótona creciente F, *F(U(.))* también es una función de utilidad válida. Esto implica que **la utilidad sigue siendo un concepto ordinal en el análisis de la conducta bajo incertidumbre**.

Por ejemplo, si *x* es una perspectiva incierta consistente de montos W_1 y W_2 con probabilidades π_1 y π_2 y las preferencias pueden ser representadas mediante la función de utilidad:

[A] $\boldsymbol{U(x)=\pi_1 \log W_1 + \pi_2 \log W_2}$

luego también será válida como función de utilidad:

[B] $\boldsymbol{V(x)=e^{U(x)}=W_1^{\pi_1} W_2^{\pi_2}}$.

Pero en este ejemplo existe una diferencia importante entre [A] y [B]. Si escribimos *u(W)=logW*, la ecuación [A] satisface la propiedad de la utilidad esperada, mientras que resulta imposible escribir [B] como el valor esperado de una función de utilidad. **En general, la propiedad de la utilidad esperada no será válida bajo transformaciones monótonas crecientes de la función de utilidad. A fin de preservar la propiedad de la utilidad esperada la transformación debe ser lineal**.

Para verificarlo, supongamos que $\boldsymbol{U=\pi_1 u\ (W_1)+\pi_2 u\ (W_2)}$ y que la sometemos a la transformación lineal $\boldsymbol{V=a+bU}$ creciente (*b>0*). Como $\pi 1+\pi 2=1$, tenemos que

[C] $\boldsymbol{V=a+ b\ (\pi_1 u\ (W_1)+\pi_2 u\ (W_2))=\pi_1(a+ b\ u\ (W_1))+\pi_2(a+ b\ u\ (W_2))}$

ecuación que sí satisface la propiedad de la utilidad esperada con la función de utilidad de Von Neumann-Morgenstern igual a *a+ b U(.)*. Es importante distinguir entre la función de utilidad elemental *U(x)* y la función de utilidad de Von Neumann-Morgenstern *V (W)*. En tanto que cualquier transformación monótona creciente de U será una función de utilidad válida para representar las preferencias por perspectivas inciertas, una transformación monótona creciente de V no producirá necesariamente una función de utilidad de Von Neumann-Morgenstern que represente a las mismas preferencias. **Las funciones de utilidad de Von Neumann-Morgenstern son únicas solamente hasta una transformación lineal**.

Por ejemplo, si las preferencias vienen representadas por la ecuación [A] previa, la función de utilidad elemental es *u (W)=log W*. Si la sometemos a una transformación monótona $v=e^{u}$ y tratamos a v como si fuera una función de utilidad de Von Neumann-Morgenstern, en tal caso las preferencias por perspectivas inciertas serían

[D] $\boldsymbol{V=\pi_1 W_1+\pi_2 W_2}$

que son distintas de [A] o [B].

A los índices únicos hasta una transformación lineal positiva se los llama índices **cardinales**. Una vez que están determinados el origen y el intervalo de los incrementos el índice cardinal queda determinado únicamente. Un ejemplo de escala cardinal es la **temperatura**. La conversión de grados Fahrenheit (F) a grados Celsius (C) viene expresada mediante

$$C = (F-32)/1.8,$$

que es obviamente lineal. La conversión inversa es F=1.8C+32.

Los índices sometidos a transformaciones lineales tienen la propiedad de que no cambia el signo de su derivada segunda. Si *W* es la riqueza y *u" (W)* es negativa (utilidad marginal decreciente), se tiene que

$$\frac{d^2}{dW^2}(a + b\,u(W)) = b\,u''(W)$$

y por consiguiente toda transformación de *u* mantendrá la propiedad de utilidad marginal decreciente. Que sea creciente o decreciente como veremos seguidamente tiene implicancias importantes sobre la actitud hacia el riesgo de un individuo. Pero ello no justifica que los cambios de la satisfacción subjetiva puedan ser comparados, dado que la función de utilidad de Von Neumann-Morgenstern sólo constituye una forma conveniente de representar las preferencias del consumidor.

Teoría de la aversión al riesgo

Luego de la axiomatización de von Neumann y Morgenstern de la hipótesis de utilidad esperada, los economistas comenzaron de inmediato a buscar las aplicaciones potenciales de la teoría a cuestiones como la elección de cartera, la actividad aseguradora, etc. Estas aplicaciones usaban modelos simples donde los resultados estaban expresados en función de un único bien, "riqueza", y por consiguiente el conjunto de resultados X era simplemente la recta real. En consecuencia, una "lotería" era definida como una variable aleatoria $z \in R$.

La conducta frente al riesgo está estrechamente vinculada con la **convexidad** de las curvas de indiferencia. Bajo certidumbre, un consumidor indiferente entre a. (2 manzanas y 0 naranjas), y b. (0 manzanas y 2 naranjas), si su curva de indiferencia es estrictamente convexa, preferirá a ambas alternativas la canasta c. (1 manzana y 1 naranja). De forma similar, como hemos visto en la teoría del consumo intertemporal, **la convexidad implica que una trayectoria suavizada de consumo es preferible a una trayectoria errática**. Analizando ahora la conducta bajo riesgo, si dibujamos un mapa de indiferencia poniendo en cada eje "ingreso en el estado 1" e "ingreso en el estado 2", vemos que **un ingreso seguro de $1 en cualquier estado es preferible a un ingreso incierto de $2 en uno de los estados y $0 en el otro**. **En otras palabras, las curvas de indiferencia convexas implican un consumidor adverso (o intolerante) al riesgo**.

Veamos ahora la vinculación entre convexidad de curvas de indiferencia y la forma de la función de utilidad elemental. Una curva de indiferencia típica de utilidad de Von Neumann–Morgenstern es la siguiente:

$$\pi_1 u(W_1) + \pi_2 u(W_2) \equiv U^0$$

Diferenciando con respecto a W_1:

$$\frac{dW_2}{dW_1} = - \frac{\pi_1 u'(W_1)}{\pi_2 u'(W_2)}$$

Con curvas de indiferencia convexas en todo su dominio, la derivada segunda:

$$\frac{d^2W_2}{dW_1^2} = - \frac{\pi_1 u''(W_1)(\pi_2 u'(W_2))^2 + \pi_2 u''(W_2)(\pi_1 u'(W_1))^2}{(\pi_2 u'(W_2))^3}$$

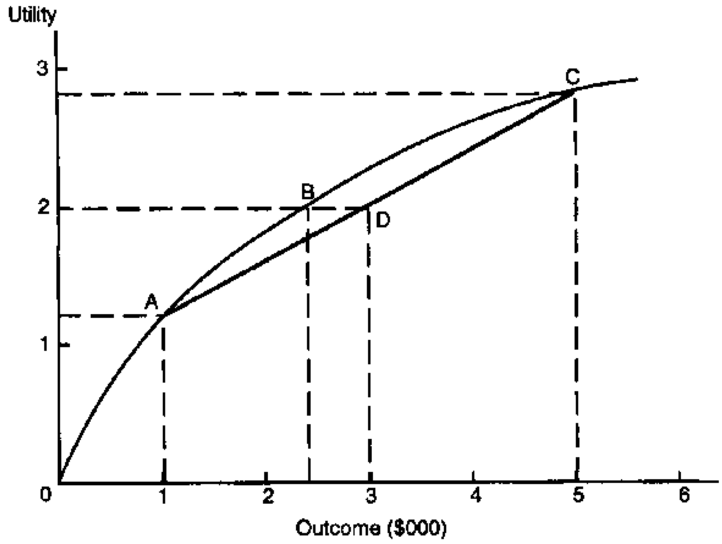
es positiva para todos los valores de W_1 y W_2. Si hacemos $W_1=W_2=W$, la derivada segunda es:

$$\frac{d^2 W_2}{dW_1^2} = - \frac{\pi_1 \pi_2 (\pi_1 + \pi_2) u''(W) u'(W_2)^2}{(\pi_2 u'(W))^3} = - \frac{\pi_1 u''(W)}{\pi_2^2 u'(W)}.$$

Esta expresión es positiva si y solamente si $u''(W) < 0$. **Luego queda demostrada la equivalencia entre el postulado de que las curvas de indiferencia son convexas por doquier y el supuesto de que la función de utilidad elemental es cóncava.**

Utilidad Esperada y Equivalente Cierto

En el gráfico adjunto, supóngase que el que toma decisiones tiene una función de utilidad como la representada en el margen. Debe elegir entre dos alternativas, una de ellas es una perspectiva segura que le produce un beneficio de $2600, y la otra una *perspectiva incierta* con una chance 50:50 de tener o bien $1000, o bien $5000. En el gráfico vemos que la perspectiva segura le depara una utilidad igual aprx. a 2.12. La utilidad de la perspectiva riesgosa viene dada por su utilidad esperada, que se calcula de la siguiente forma:



U ($1000) (0.5) + U ($5000) (0.5).

Según vemos, en el gráfico está representada esta utilidad como 1.2 (0.5) + 2.8 (0.5) = 2. Al resultar 2.12 > 2, la decisión de aceptar el retorno seguro de $2600 será preferida a la decisión con riesgo. Esto sucede aunque el valor monetario esperado de la alternativa riesgosa (VME = $1000. (0.5)+$5000. (0.5)) = $ 3000 sea superior al valor monetario de la alternativa cierta ($2600).

Esta figura ilustra los conceptos de aversión al riesgo y de equivalente cierto. La aversión al riesgo de la persona resulta evidente por el hecho de que, aunque la perspectiva riesgosa tenga un valor monetario esperado de $3000, no será preferida a la alternativa segura, pues su valor de utilidad de 2 (por ser riesgosa) es inferior a la utilidad de la perspectiva segura con un valor de 2.3. Véase que asociado al punto B de la curva de utilidad, el importe monetario de la alternativa riesgosa corresponde a un monto de $2400. En otras palabras, para esta persona, la perspectiva riesgosa es equivalente, en términos de utilidad, a la percepción de $2400. Este monto de $2400 es referido entonces como el **equivalente cierto de la perspectiva riesgosa**. En general, la teoría de la utilidad esperada implica que, frente a cualquier opción riesgosa a la que haga frente esta persona, siempre será posible especificar un valor seguro (=cierto) que tenga el mismo valor que la perspectiva riesgosa. O sea, en otras palabras, que la persona estará indiferente entre tomar la opción riesgosa y recibir el importe seguro así determinado. A este importe seguro se lo conoce como **equivalente cierto (EC)** de la opción riesgosa. También se tiene que, como diferentes personas tendrán funciones de utilidad diferentes, también tendrán distintos estimados EC de una misma opción riesgosa.

La diferencia entre el valor medio o esperado de un proyecto riesgoso (VE) y su equivalente cierto EC (es decir, VE – EC en el caso de la perspectiva monetaria) constituye la **prima de riesgo** del proyecto riesgoso. En nuestro ejemplo, la prima está dada por la distancia BD, que implica una prima igual a $3000 - $2400 = $600: esta persona está preparada para dar $600 "esperados" a fin de evitar el proyecto riesgoso. Naturalmente, otra persona podría tener una prima al riesgo diferente para el mismo proyecto.

Si el que toma decisiones es **neutro al riesgo**, su utilidad será una función lineal, lo que implica una prima de riesgo nula – que es lo que sucedería si la línea ADC de la figura fuera la función de utilidad relevante. A la inversa de un intolerante al riesgo, si es **tolerante del riesgo** (o amante del riesgo) se tendrá: (1) una función de utilidad convexa, con pendiente positiva y creciente que refleja una utilidad marginal creciente, y (2) una prima al riesgo negativa, dado que la persona estaría dispuesta a **pagar una prima para afrontar la oportunidad riesgosa**.

Los resultados obtenidos están incorporados en el cuadro siguiente.

		Atributo	
Forma de la función de utilidad *U(X)*	**Actitud hacia el riesgo**	**Prima de riesgo *(VE - EC)***	**Utilidad marginal *(dU/dX>0)***
Cóncava	aversión *(VE > EC)*	positiva	decreciente $(d^2U/dX^2 < 0)$
Lineal	Neutralidad *(VE = EC)*	cero	constante $(d^2U/dX^2 = 0)$
Convexa	preferencia *(VE < EC)*	negativa	creciente $(d^2U/dX^2 > 0)$

Medidas de Aversión al riesgo

Medida de aversión absoluta al riesgo Hemos visto que la medida de aversión al riesgo está estrechamente relacionada con la concavidad de la función de utilidad del individuo, es decir, a mayor concavidad de la función de utilidad, tanto más adverso al riesgo será el individuo. Arrow y Pratt[2] han propuesto una medida de aversión al riesgo invariante a cambios de la función de utilidad esperada mediante una transformación afín, normalizando la segunda derivada de la función por la primera derivada. Así obtenemos la medida de Arrow-Pratt de aversión (absoluta) al riesgo:

$$R_A(X) = \frac{U''(X)}{U'(X)}$$

Como caso interesante, se tiene que la función de utilidad exponencial $U(X)=-e^{-aX}$ implica una medida constante de aversión al riesgo $R_A(X)$, igual al parámetro a.

Costo monetario del riesgo de Arrow-Pratt Definiendo como costo monetario del riesgo al *monto de riqueza cierta que el consumidor estaría dispuesto a ceder a fin de evitar el riesgo considerado*, y denotándolo como ζ, se satisfará la ecuación siguiente:

$$\zeta = -\frac{U''(E(X))}{2U'(E(X))}\sigma_X^2$$

En esta expresión, σ_X^2 es igual a $(X - E(X))^2$, es decir, la varianza del evento considerado. Luego, el costo del riesgo viene dado por el producto de una medida subjetiva (el coeficiente de aversión al riesgo ya visto) y otra medida objetiva como la varianza de su riqueza.

Medida de aversión al riesgo relativa Sea $V(W)$ la función de utilidad indirecta de un agente, donde W = riqueza monetaria y supongamos que dicho agente maximiza el valor esperado de la utilidad indirecta de su riqueza según el enfoque de Von Neumann y Morgenstern de comportamiento en condiciones de riesgo. La variable W es una v.a. que, para simplificar, asume n valores discretos X_i con probabilidades Q_i no negativas que suman la unidad. Por lo tanto, resuelve:

$$\text{Máximo } E[V(W)] = \sum_{i=1}^{n} Q_i \; V(W_i).$$

El **coeficiente de aversión relativa al riesgo de Arrow-Pratt** (ρ) es la elasticidad de la función derivada de utilidad indirecta, $V'(W)$, es decir $\boldsymbol{\rho = - V''(W)\, W / V'(W)}$ Luego se tiene:

$$\frac{\zeta}{E(W)} = \tfrac{1}{2}\rho \left(\frac{\sigma(W)}{E(W)}\right)^2$$

ζ es el costo monetario absoluto del riesgo. Ésta es una medida muy útil, con un significado intuitivamente preciso: el costo monetario absoluto del riesgo, en proporción a la riqueza esperada, está directamente relacionado con el *coeficiente de variación de la riqueza*, que es una medida de la dispersión de la distribución de probabilidades de la riqueza, elevado al cuadrado. Éste es un parámetro objetivo. El otro parámetro interviniente es el *coeficiente de aversión relativa al riesgo de Arrow-Pratt ρ*. En problemas de elección intertemporal, la elasticidad de sustitución intertemporal a menudo es la misma que el coeficiente de aversión relativa al riesgo.

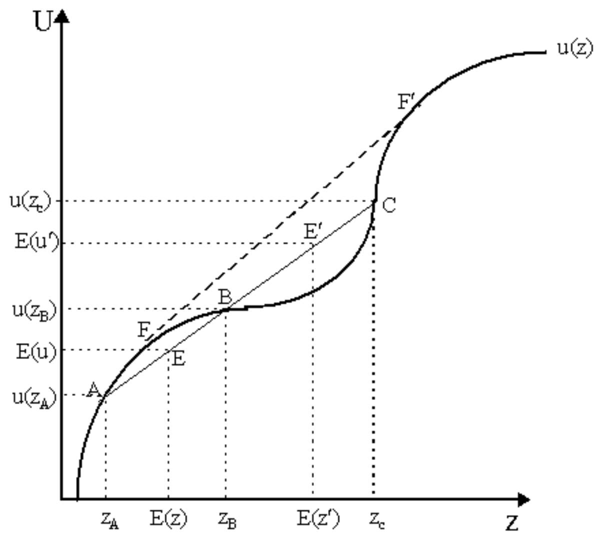
La evidencia experimental y empírica es más bien consistente con la hipótesis de aversión absoluta decreciente al riesgo.

---

[2] John Pratt, Risk Aversion in the Small and in the Large, Econometrica, 1964.

Apuestas, seguro y diversificación

Si no hay otras restricciones sobre la forma de la función de utilidad, la hipótesis de la utilidad esperada es consistente tanto con una conducta inclinada a asumir riesgos como con una conducta destinada a evitarlos. Friedman y Savage sostuvieron que si la función de utilidad tiene la forma indicada en esta página, **un individuo comprará seguro y billetes de lotería al mismo tiempo**. Observen que *u(z)* es cóncava hasta el punto de inflexión B y luego es convexa hasta el nuevo punto de inflexión C, después del cual pasa a ser cóncava nuevamente.



Friedman y Savage trataron de usar esta construcción para explicar por qué la gente puede tomar riesgos de baja probabilidad pero con pagos elevados (p.ej. billetes de lotería), y al mismo tiempo asegurarse contra riesgos moderados con pagos moderados (p.ej. comprar seguro con el pasaje aéreo). Para apreciar esta posibilidad, supongan que se encuentran en B, punto de inflexión entre la aversión al riesgo y la propensión al riesgo. Supongan que tienen frente a ustedes dos loterías, una que otorga A o B y otra que otorga B o C. Estas loterías están representadas por las cuerdas entre sus pagos respectivos AB y BC. La utilidad esperada de la primera se indica como *E (z)* y está representada en el punto E, en el cual *E (u)* resulta obviamente menor que la utilidad del resultado esperado de la primera apuesta, *u (E (z))*. Luego un agente adverso al riesgo pagará una prima por evitarlo. La segunda apuesta da lugar a una utilidad esperada *E (u)* en el punto E' de la cuerda BC – mayor que la utilidad del resultado esperado *u (E (z'))* y, en este caso, un apostador amante del riesgo pagaría una prima para tomarla. De tal forma puede considerarse que la aversión al riesgo en el tramo AB es un caso de asegurarse contra pequeñas pérdidas, en tanto que la inclinación al riesgo en BC es un caso de compra de billetes de lotería.

Markowitz ha controvertido la hipótesis de que la gente, al menos a nivel agregado, tenga curvas de utilidad con tales dobles inflexiones. Notó que una persona en F también aceptaría una apuesta que lo dejase en F'. Recíprocamente, una persona en F' o levemente debajo de ese punto no pagaría una prima para no terminar en F, es decir, no se asegurará contra situaciones de grandes pérdidas de baja probabilidad. Finalmente, los muy ricos, arriba de F' nunca entrarán en apuestas justas – algo que entra en contradicción con fenómenos empíricos tales como los casinos de Monte Carlo, etc. Lo que propuso en su lugar fue que los z sean considerados no como "niveles de ingreso" como Friedman y Savage, sino como "cambios del ingreso". Agregó asimismo un punto adicional de inflexión en la región inferior. El ingreso "normal" de la gente – cualquiera fuera su ubicación en la pirámide distributiva – vendría dado por un punto como B y lo demás reflejaría desvíos de este ingreso promedio. De esta forma, la paradoja aparente de la lotería–seguro sería resuelta sin invocar las extrañas implicancias de la hipótesis original de Friedman–Savage.