Тема 1. Введение

1. *Теория фон Неймана-Моргенштерна*

Простой лотереей называется любой *конечный* набор упорядоченных пар чисел вида $[x_1,p_1;\ldots;x_n,p_n]$, где x_i – денежные выигрыши, а p_i – вероятности получения каждого из них, $\sum p_i=1$. Введем следующие аксиомы на множестве P всех простых лотерей:

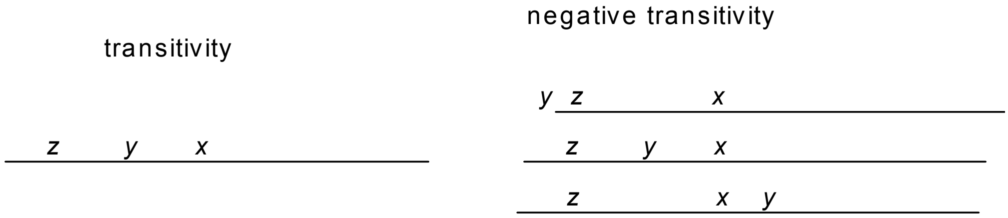
- A1 **Слабая упорядоченность** строгого предпочтения ($\succ$ или $\prec$)

A) **Асимметрия** $p\succ q \Rightarrow \neg(q\succ p)$, $p,q\in P$;

Б) **Негативная транзитивность** $\neg(p\succ q)\&\neg(q\succ r) \Rightarrow \neg(p\succ r)$, или $p\succ r \Rightarrow p\succ q \vee q\succ r$, $\forall p,q,r\in P$ (эквивалентная форма – из $p\succ r \Rightarrow p\succ q\vee q\succ r$ следует $\neg(p\succ q\vee q\succ r) \equiv \neg(p\succ q)\ \&\ \neg(q\succ r) \Rightarrow \neg(p\succ r)$).

A) означает **иррефлексивность** $\neg(p\succ p)$ for all $p\in P$ (достаточно взять $q=p$)

Б) слабее **транзитивности** $(p\succ q)\&\ (q\succ r) \Rightarrow (p\succ r)$, т.е. подразумевает ее: все транзитивные отношения негативно транзитивны, но обратное неверно:



В силу свойства A1.Б) определим $p\sim q$ как $\neg(p\succ q)\&\neg(q\succ p)$. С учетом этого определения, A1.A) означает также **полноту** предпочтений: для каждых двух простых лотерей имеем $p\succ q \vee q\succ p \vee p\sim q$. Отсюда же можно определить **нестрогое предпочтение** $p\succcurlyeq q$ как $p\succ q \vee p\sim q$, которое также можно взять за базовое отношение. Из этого последнего, в свою очередь, можно определить строгое предпочтение (как $p\succ q \equiv p\succcurlyeq q\ \&\ \neg(q\succcurlyeq p)$) и безразличие (как $p\sim q \equiv p\succcurlyeq q\ \&\ q\succcurlyeq p$). Тогда из определения $\succ$ и полноты $\succcurlyeq$ ($q\succcurlyeq p \vee p\succcurlyeq q$ для всех $p,q\in P$) будет следовать асимметрия $\succ$, а из транзитивности нестрогого предпочтения – негативная транзитивность строгого (в первой форме). Таким образом, за аксиому A1 можно взять асимметрию и негативную транзитивность отношения $\succ$, или полноту и транзитивность отношения $\succcurlyeq$.

- A2 **Независимость**

$\forall p,q,r\in P\ \forall\alpha\in[0,1]\ p\succ q \Rightarrow \alpha p+(1-\alpha)r\succ\alpha q+(1-\alpha)r$

Сильное свойство, требующее инвариантности предпочтений при любых линейных (вероятностных) комбинациях лотерей. Экспериментальные опровержения этой аксиомы известны давно (см. ниже), и большинство современных теорий выбора в условиях риска построены прежде всего на ослаблении A2.

- A3 **Непрерывность**

$\forall p\succ q$ и $q\succ r\in P$, $\exists\alpha,\beta\in(0,1)$ т.ч. $\alpha p+(1-\alpha)r\succ q$ и $q\succ\beta p+(1-\beta)r$. В несколько иной формулировке, это условие звучит так: для всех $p\succ q$ и $q\succ r\in P$, существуют такие открытые множества чисел α и $\beta\in(0,1)$, что $\alpha p+(1-\alpha)r\succ q$ и $q\succ\beta p+(1-\beta)r$.

Стандартная аксиома, также небесспорная с содержательной точки зрения. Понятно, что \$1$\succ$\$0.01$\succ$смертной казни, но верно ли, что существует $\alpha<1$ т.ч. α\$1+(1-α)смертная казнь $\succ$ \$0.01? «Теоретически» этот пример опровергает A3, однако «практически» кто из нас не перебегал хотя бы раз улицу перед близко идущим транспортом, подвергая свою жизнь опасности смерти с ненулевой вероятностью во имя несопоставимо малого выигрыша?

A1-A3 гарантируют существование функции полезности Неймана-Моргенштерна, которая является основной парадигмой теории экономического выбора в условиях риска, и в частности теории игр.

Кроме того, поскольку теория выбора в условиях риска наделяется дескриптивной функцией («даже если в действительности люди не считают матожиданий каких-либо функций, они ведут себя так как если бы они их максимизировали») – эти аксиомы играют также роль постулатов рационального поведения в том смысле, что любой рациональный человек должен их соблюдать.

Теорема (Нейман-Моргенштерн, 1944; Йенсен, 1967; Фишберн, 1970 [1974]):
Пусть выполнены условия А1-А3 на множестве всех простых лотерей. Тогда на множестве *исходов* существует такая функция полезности $u(x)$, что предпочтения индивидов на множестве *лотерей* представимы в виде математического ожидания этой функции, $U(p)=\sum p(x)u(x)$, и это представление обладает следующими свойствами:

1. Сохранение порядка $p \succ q \Leftrightarrow U(p)=\sum p(x)u(x) > \sum q(x)u(x)=U(q)$ for all $p, q \in P$.
2. Линейность по вероятностям $\forall p,q,r \in P$ $u[\alpha p+(1-\alpha)q] \sim \alpha u(p)+ (1-\alpha)u(q)$
3. Определенность с точн. до афинной трансформации $V(p)=a+bU(p)$, $\forall a, b>0$

Сохранение порядка гарантирует представление предпочтений при помощи действительнозначной функции $U(.)$, что позволяет трактовать ее как функцию полезности. Линейность по вероятностям на всех *простых* лотереях обеспечивает представление любой такой функции как математическое ожидание полезностей составляющих ее лотерей (интерпретируя достоверный исход как лотерею [x,1]). Именно это свойство обеспечивает представление предпочтений в виде ожидаемой полезности. Наконец, определенность с точностью до афинной (выпуклой и вогнутой одновременно) трансформации характеризует эту полезность как *кардинальную для класса функций* – частичный возврат к кардиналистским представлениям о полезности после ординалистской революции Хикса-Аллена в середине 1930-х.

Доказательство: Проведем доказательство в три этапа. На первом мы докажем ряд предварительных лемм:

(1). $p \succ q$, $\forall \alpha \in (0,1) \Rightarrow p \succ \alpha p + (1-\alpha)q \succ q$

(2). $p \succ q$, $\forall \beta,\alpha \in (0,1)$: $\beta>\alpha \Rightarrow \beta p+(1-\beta)q \succ \alpha p + (1-\alpha)q$

(3). $p \succcurlyeq q \succcurlyeq r$, $p \succ r \Rightarrow \exists$ единственное λ: $q \sim \lambda p +(1-\lambda)r$

Докажем эти леммы.

(1). $p = \alpha p + (1-\alpha)p \succ \alpha p + (1-\alpha)q \succ \alpha q + (1-\alpha)q = q$ (оба предпочтения следуют из А2, примененной последовательно к разным слагаемым)

(2). определим $\gamma=(\beta-\alpha)/(1-\alpha)$ – заметим, что из определений β и α всегда будет $\gamma \in [0,1]$. Заметим, что $\beta p+(1-\beta)q$ можно переписать в виде $[(1-\alpha)\beta p+(1-\alpha)(1-\beta)q]/(1-\alpha) = [(\beta-\alpha+\alpha-\alpha\beta)p+(1-\alpha-\beta+\alpha\beta)q]/(1-\alpha) = [(\beta-\alpha)p + (1-\beta+\alpha-\alpha)\alpha p]/(1-\alpha) + [(1-\beta)q - (1-\beta)\alpha q]/(1-\alpha) = [(\beta-\alpha)/(1-\alpha)]p + [(1-\alpha-\beta+\alpha)/(1-\alpha)]\alpha p + [(1-\alpha-\beta+\alpha)/(1-\alpha)]q - [(1-\alpha-\beta+\alpha)/(1-\alpha)]\alpha q = \gamma p + (1-\gamma)[\alpha p + (1-\alpha)q]$, где $(1-\alpha-\beta+\alpha)/(1-\alpha)=1-\gamma$. Обозначим теперь $\alpha p + (1-\alpha)q = r=s$. В силу (1), $p \succ \alpha p + (1-\alpha)q \equiv s$. Взяв γ в качестве веса, можно записать $\gamma p + (1-\gamma)s \succ \gamma s + (1-\gamma)r$ по А2 (т.к. $r=s$), что эквивалентно $\gamma p + (1-\gamma)[\alpha p + (1-\alpha)q] \succ [\alpha p + (1-\alpha)q] \Leftrightarrow \beta p + (1-\beta)q \succ \alpha p + (1-\alpha)q$, т.к. $\gamma+\alpha-\alpha\gamma=\beta$.

(3). Если $p \sim q$ или $q \sim r$, то годится $\lambda=1$ и $\lambda=0$, соотв. Если $p \succ q \succ r$, то определим $\lambda \equiv \sup\{\alpha \in (0,1): q \succ \alpha p+(1-\alpha)r\}$, где указанное множество непусто, т.к. приложимо к $\alpha=0$. Для всех $\alpha \in (\lambda,1]$ имеем $\alpha p+(1-\alpha)r \succ q$ по определению, а для всех $\alpha \in [0,\lambda)$ - $q \succ \alpha p+(1-\alpha)r$ в силу (2), причем множества всех таких α в обоих случаях открыты в силу непрерывности А3. Остается показать, что таким образом определенное λ обеспечивает безразличие между $\lambda p+(1-\lambda)r$ и q. Пусть это не так, и $\lambda p+(1-\lambda)r \succ q \succ r$ (т.е. q лежит между лотереей с λ и худшей лотереей r). Тогда в силу А3 найдется такое $\alpha \in (0,1)$, что $\alpha[\lambda p+(1-\lambda)r]+(1-\alpha)r \succ q \Leftrightarrow \alpha\lambda p+(1-\lambda\alpha)r \succ q$ – но это последнее утверждение несовместимо с определением λ, ибо $\lambda>\alpha\lambda$, а для всех чисел меньших чем λ лотерея q должна была быть лучше чем $\alpha p+(1-\alpha)r$. Полученное противоречие доказывает, что $q \succcurlyeq \lambda p+(1-\lambda)r$. Заменяя $\succ$ на $\prec$ и повторяя рассуждение, исключаем и $q \prec \lambda p+(1-\lambda)r$, так что остается безразличие. Единственность следует из (2): для любого иного λ' будем иметь $\succ$.

На втором этапе покажем, что A1-A3 гарантируют существование функции полезности, представимой в виде линейной комбинации полезностей исходов с вероятностными весами, т.е. в виде ожидаемого значения этой функции, который может служить индексом полезности лотерей. Поскольку леммы (1)-(3) справедливы для всех лотерей, они применимы и к максимальному выигрышу M и минимальному выигрышу m. Тогда в силу (3) для любой лотереи p существует единственное $u(p)$: $p \sim u(p)M + (1-u(p))m$ (где для M и m имеем соотв. $u(M) =1$ и $u(m) =0$). Для любых двух лотерей p, q справедливо одно из трех соотношений предпочтений (в силу полноты), так что из (2) будем иметь $u(p) >(<,=)\ u(q) \Leftrightarrow p \succ (\prec, \sim)\ q$. Это показывает, что определенный таким образом индекс $u(\cdot)$ сохраняет предпочтения, т.е. является индексом полезности. Для доказательства линейности по вероятностям рассмотрим лотерею $\lambda p+(1-\lambda)q \sim \lambda[u(p)M + (1-u(p))m] + (1-\lambda)[u(q)M + (1-u(q))m] = M[\lambda u(p)+(1-\lambda)u(q)] + m[\lambda-\lambda u(p)+1-\lambda-u(q)+\lambda u(q)] = M[\lambda u(p)+(1-\lambda)u(q)] + m[1-\lambda u(p)-(1-\lambda)u(q)]$. Определяя теперь полезности всех лотерей, получаем $u(\lambda p+(1-\lambda)q) = u(M)[\lambda u(p)+(1-\lambda)u(q)] + u(m)[1-\lambda u(p)-(1-\lambda)u(q)] = \lambda u(p)+(1-\lambda)u(q)$.
Докажем наконец, что функция полезности определена с точностью до афинной трансформации, т.е. что $\sum u(x)q(x) > \sum u(x)s(x) \Leftrightarrow \sum v(x)q(x) > \sum v(x)s(x) \Leftrightarrow v(x)= u(x)a+b$, $a>0$. Если представления по u и по v одинаковы, то $\sum u(x)q(x)$. Рассмотрим сперва ожидаемую полезность функции полезности v при том что она и функция u связаны линейно. Для любого p имеем $E[v(p)]=E[au(p)+b] = [\sum u(x)p(x)]a+b = \sum p(x)[u(x)a+b] = \sum v(x)p(x)$, так что оба представления полезности идентичны.
Напротив, предположим что представления по u и по v однозначны. Без ограничения общности примем $p \succ q \succ r$ и определим $u(q) = \alpha u(p) + (1-\alpha)u(r)$, так что $q \sim \alpha p + (1-\alpha)r$. Из этого выражения получаем, что $\alpha = [u(q)-u(r)]/[u(p)-u(r)]$. Но поскольку по условию и v представляет те же предпочтения, то имеем также и $v(q) = \alpha v(p) + (1-\alpha)v(r) = \alpha[v(p) - v(r)] +v(r)$. Подставив сюда полученное ранее выражение для α, получаем

$$v(q) = \frac{u(q)-u(r)}{u(p)-u(r)}\left(v(p)-v(r)\right)+v(r) = \frac{v(p)-v(r)}{u(p)-u(r)}u(q)+v(r)-u(r)\frac{v(p)-v(r)}{u(p)-u(r)} = au(q)+b$$

Теорема полностью доказана.

Теория Неймана-Моргенштерна была и остается фундаментальным инструментом не только теории игр, но и всей теории выбора в условиях риска. Тем не менее уже вскоре после ее появления (в начале 1950-х) были получены, и затем подтверждены многочисленные экспериментальные данные (так называемые «парадоксы рациональности»), которые опровергают аксиомы A1-A3, делая тем самым невозможным существование функции ожидаемой полезности в виде Неймана-Моргенштерна. Первым и до сих пор самым знаменитым контрпримером остаются парадоксы Алле (Allais, 1951)

Парадокс Алле (версия Kahneman and Tversky 1979).

Пара 1:
A: 250, 0.33; 240, 0.66; 0, 0.01 B: 240, 1.0
Пара 2:
C: 250, .33; 0, 0.67 D: 240, .34; 0, 0.66

88% выбирают B в первой паре, и 83% выбирают C во второй – но комбинации B-C и A-D нарушают A2, т.к. $u(240)>.33(250)+.66u(2400)+.01u(0)$ или $.34u(240)>.33u(250)+.01u(0)$, в противоречии с C>D, или $.33u(250)>.34u(240)$. Покажем, что это явление означает нарушение A2: Определим $\delta=z=[240,1]$, $x=[0,1]$ and $Y=[250, 33/34; 0, 1/34]$. В силу полноты предпочтений δ может быть лучше или хуже или равно Y; по A2, таким же должно быть предпочтение в отношении $\lambda\delta$ и λY, дополненных одной и той же лотереей с вероятностью $(1-\lambda)$. При $\lambda=0.34$ (и $1-\lambda=0.66$), получаем
Пара 1:

$\lambda Y+(1-\lambda)z$ = .34*(250*33/34+0*1/34)+.66*240 = 250*(33*34)/(34*100) + 0*(1*34)/(34*100) + 240*.66 = 250*.33 + 240*.66 + 0*.01 = A

$\lambda\delta+(1-\lambda)z$ = .34*240+.66*240 = 240*1 = B
Пара 2:
λY+(1-λ)x = .34*(250*33/34+0*1/34)+.66*0 = 250*(33*34)/(34/100) + 0*(.01+66) = 250*.33 + 0*.67 = C
$\lambda\delta+(1-\lambda)x$ = .34*240 + .66*0 = D
Это явление называется эффектом общих последствий – добавление одного и того же z или x небезразлично (**common consequence effect, CC**).
Другой пример парадокса Алле – эффект общего отношения (вероятностей в двух лотереях) (**common ratio effect, CR**):

Пара 1:

E: [240, 1; 0,0]	F: [250, 0.8; 0, 0.2]

Пара 2:

G: [240, 0.25; 0, 0.75]	H: [250, 0.2; 0, 0.8]

Положим λ=0.25, p=E, q=F, r=0, сразу заметим, что требование A2 означает, что $E=p>F=q$ если и только если $\lambda p+(1-\lambda)r < \lambda q+(1-\lambda)r$, или .25*(240*1)+.75*0 < .25*(250*.8+0*.2)+.75*0, or 240*.25+0*.75 < 250*.2+0*.8 т.е. G < H,- но этого не наблюдается. Эти эффекты представлены на так называемом треугольнике Маршака--Машины CC - квадраты, CR - круги.

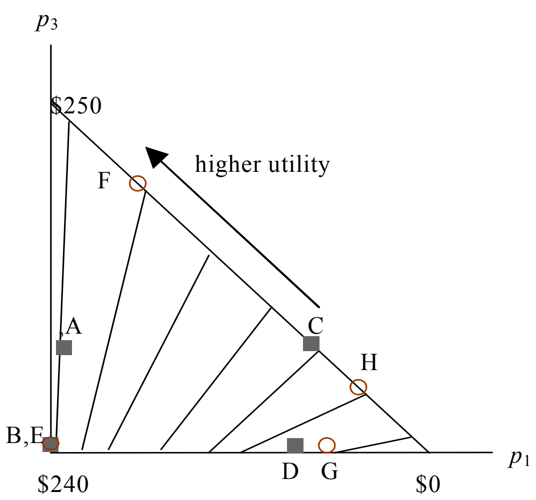


Рис.1 Парадокс Алле
Как нетрудно заметить, полезность НМ означает, что кривые безразличия (все лотереи, столь же желательные, что и $X=[a, p_3; b, p_2; c, p_1]$, $a\succ b\succ c$) на Рис.1 должны быть параллельны, ибо

$U(X)=u(a)p_3 + u(b)(1-p_3-p_1) + u(c)p_1 \Rightarrow$
$p_3=[U(X)-u(b) + p_1(u(b)-u(c)]/[u(a)-u(b)]$
Парадоксы Алле означают нарушение этой линейности – кривые безразличия должны иметь разный наклон.

Несмотря на эти и другие парадоксы, теория ожидаемой полезности (ОП) Неймана-Моргенштерна лежит в основе практически всех экономических моделей выбора в условиях риска – факт сам по себе примечательный!