Explain what a certainty equivalent is and how to calculate it for a given lottery [p, A; 1-p, B] and a given utility function u().

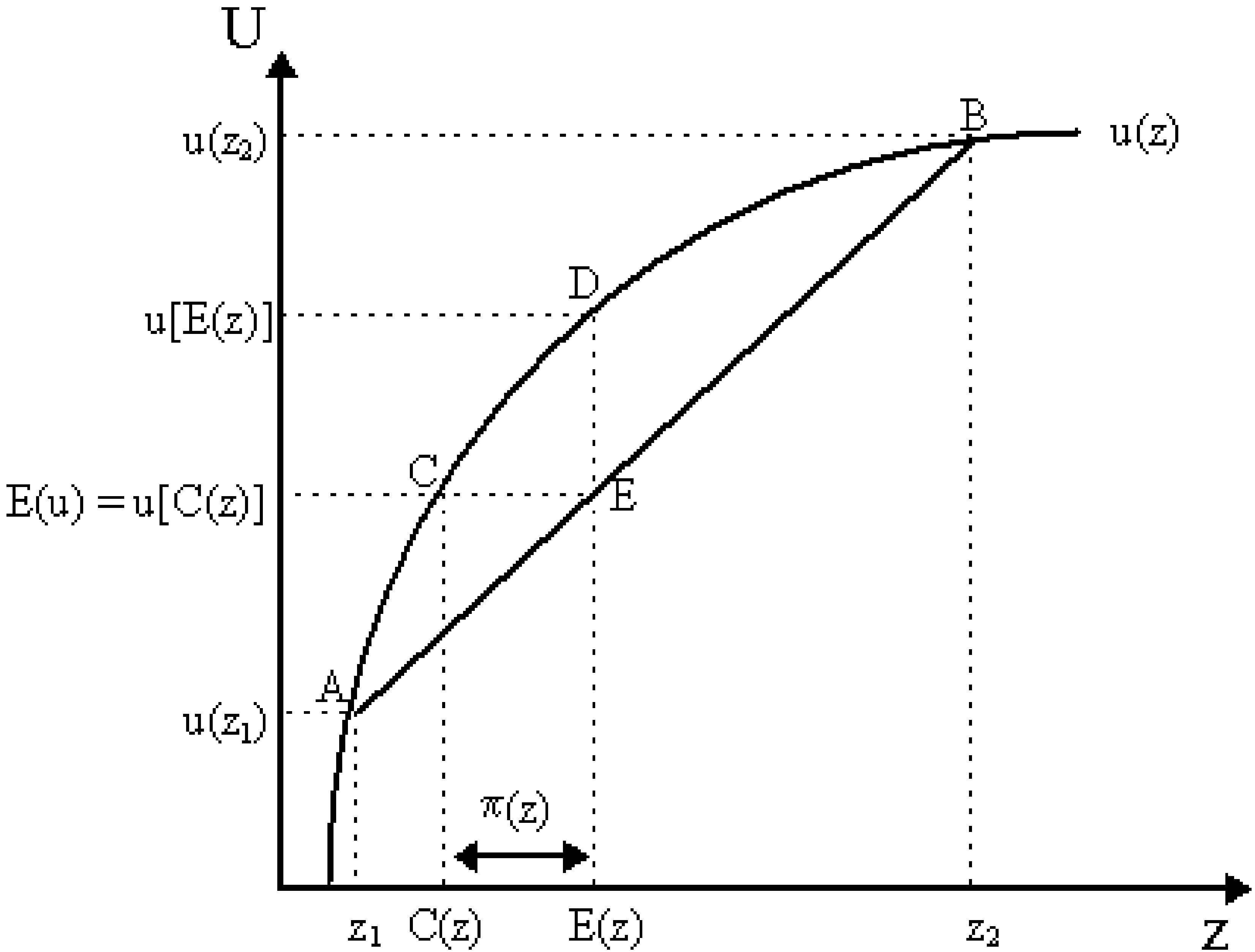
Certainty Equivalent: The amount of payoff that an agent would have to receive to be indifferent between that payoff and a given gamble is called that gamble's 'certainty equivalent'.

To Calculate the Certainty Equivalent:

Let z1 be the payoff for A and z2 be the payoff for B

Lets assume the utility function to be concave.

The expected payoff will be $p*z1 + (1-p)*z2$ and
the expected utility will be $p*u(z1) + (1-p)*u(z2)$



U
u(z₂)
u[E(z)]
E(u) = u[C(z)]
u(z₁)
B
u(z)
D
C
E
A
π(z)
z₁
C(z)
E(z)
z₂
Z

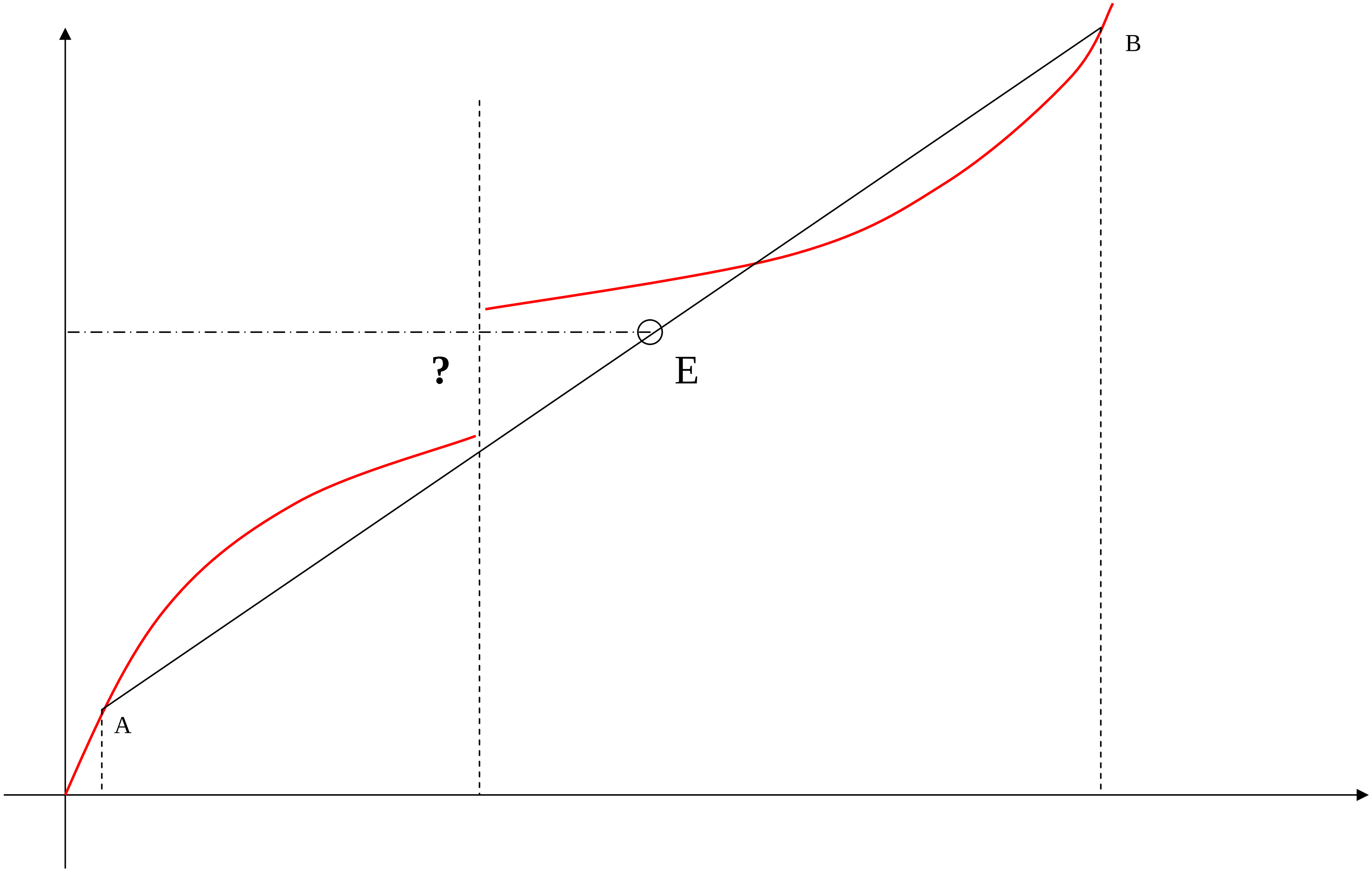
The continuity property is often chosen as one of the axioms of utility theory: If A, B and C are arbitrary outcomes, then A>B>C implies that, for any decision maker, there exists a probability p such that the decision maker is indifferent between the lottery [p,A; 1-p,C] and the outcome B. Why is this not necessarily true?

Even if the utility and maximum expected utility principles continue to hold, it is not necessary the case that one can use the certainty equivalents of lotteries to compare them. Why not?

According to continuity property if A > B > C then there always exists a positive p such that,

$$p*A + (1-p)* C = B$$

Let us consider the following utility function:



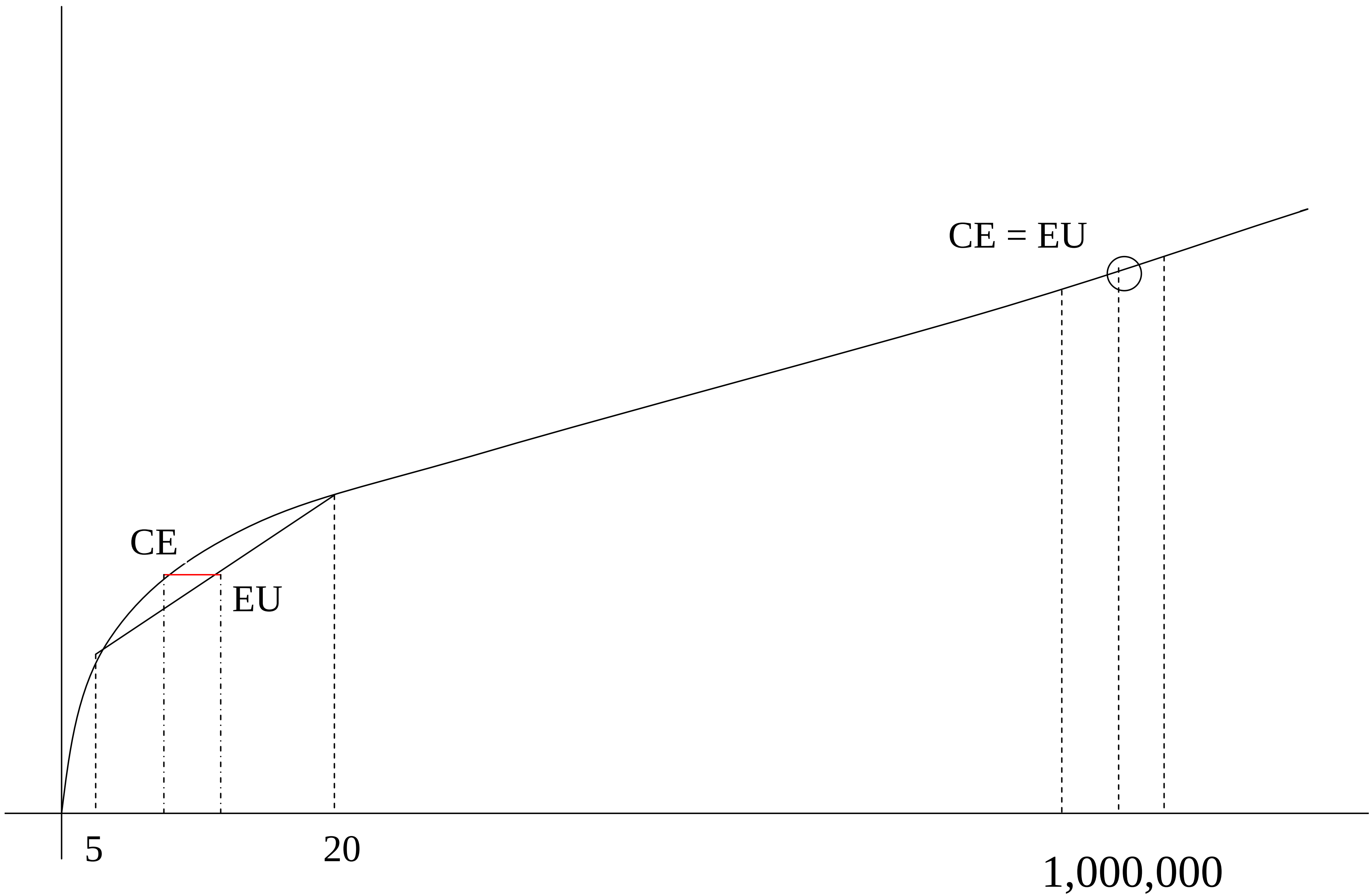
Give an example that demonstrates that two decision makers with the same utility function but different wealth can make different decisions.

Consider the following lottery:

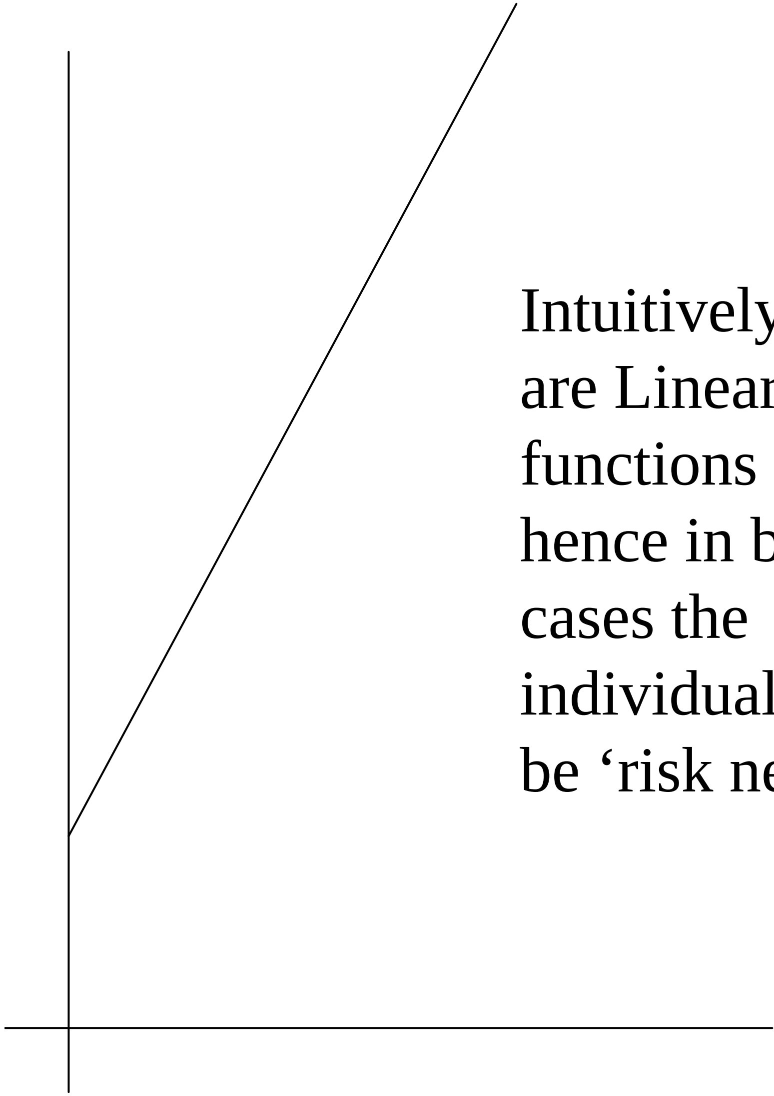
You can either win \$10 with a probability of 0.5 or lose \$5 with a probability of 0.5.

You can either play the lottery or
leave the game for an amount of \$5.

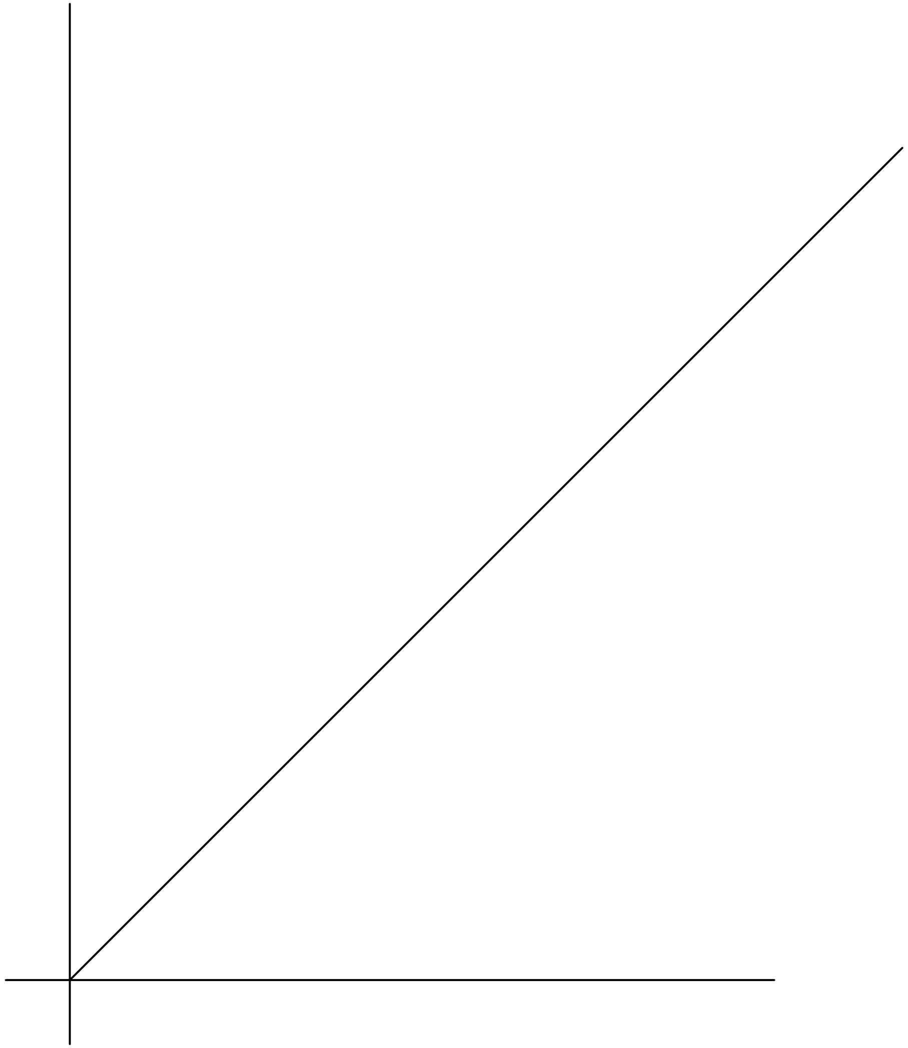
Now consider 2 individuals with initial wealth \$10 and \$1,000,000 but with the same utility function.


CE = EU
CE
EU
5
20
1,000,000

Explain in detail why the utility functions u(reward) = 3 reward + 4 and u(reward) = reward are considered to be equivalent.



Intuitively: Both are Linear functions and hence in both cases the individual will be 'risk neutral'



Initially = $5.

Can win $5 with 0.5 probability and lose $5 with 0.5 probability.

Alternatively just not play the game.

For: U(reward) = 3*reward + 4

U ($5) = 3 * 5 + 4 = 19

U ($10) = 34

U ($0) = 4

Hence Expected Utility = 34*0.5 + 4 * 0.5 = 19

Thus the individual will be indifferent to both choices

For: U(reward) = reward

U ($5) = 5

U ($10) = 10

U ($0) = 0

Hence Expected Utility = 10*0.5 + 0 * 0.5 = 5

Thus again individual will be indifferent to both choices.

Hence we can see that the decision taken by both the curves is the same.

Derive the a) the Bellman equations and b) the update rules of value iteration for maximizing the expected total undiscounted utility for a given Markov decision problem and the concave exponential utility function u(reward) = $-gamma^{reward}$ where 0 < gamma < 1 is a given constant.

$U(s) = \max_a [\ r(s,a) + \text{sum}_{s'} \{\ T(s,a,s') * U(s')\}\ ]$

Thus,

$U(s) = \max_a [\ \text{sum}_{s'} \{\ T(s,a,s') * \{\ r(s, a) + U(s')\ \}\ \}]$ if S is not a goal state

$= 0$ if S is a goal state

We have $U(r) = -\ gamma^{r}$

Hence, $r = \log_{gamma}(-U)$

Thus $U^{-1}(x) = \log_{gamma}(-U)$

$U(s) = \max_a [\, \text{sum}_{s'} \{ T(s, a, s')\, U(r(s,a,) + U^{-1}(U(s'))) \}]$

$U(s) = -\max_a [\, \text{sum}_{s'} \{ T(s, a, s')\ \text{gamma}^{r(s,a) + \log_{\text{gamma}} (-U(s'))} \}]$

$U(s) = -\max_a [\, \text{sum}_{s'} \{ T(s, a, s')\ \text{gamma}^{r(s,a)} * (-U(s')) \}]$

$U(s) = \max_a [\, \text{sum}_{s'} \{ T(s, a, s')\ \text{gamma}^{r(s,a)} * U(s') \}]$

if s is not a goal state

$U(s) = U(0) = -\text{gamma}^0$ if s is a goal state

$= -1$