26: Utility of Money, Simulating Probabilities, Jensen's Inequality

Lisa Yan

June 5, 2020

Quick slide reference

LIVE

Simulating Probabilities

from scipy import stats

stats.uniform , ...

generate ~~Bea~~ Binomial

`random.random()` function

Since computers are deterministic, **true** randomness does not exist.

We settle for pseudo-randomness: A sequence that looks random, but is actually deterministically generated.

`random.random()`, `np.random.random()`

- returns a float uniformly in [0.0, 1.0) with the Mersenne Twister:
- 53-bit precision floating point, repeats after 2**19937-1 numbers
- Seed number: X_0 used to generate sequence $X_1, X_2, \ldots, X_n, \ldots$

Initialization [edit]

The state needed for a Mersenne Twister implementation is an array of *n* values of *w* bits each. To initialize the array, a *w*-bit seed value is used to supply x_0 through x_{n-1} by setting x_0 to the seed value and thereafter setting

$$x_i = f \times (x_{i-1} \oplus (x_{i-1} >> (w-2))) + i$$

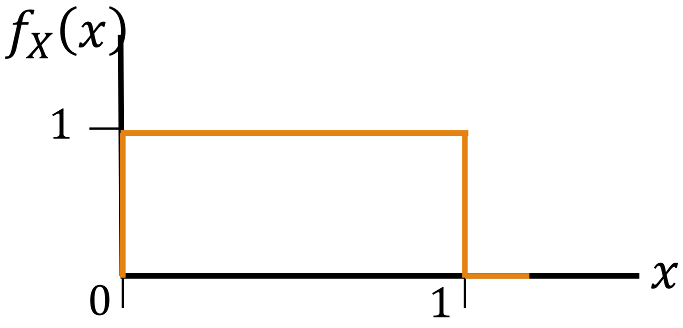
pset1_code — vim cs109_pset1.py — 73×18

```python
def q14(seed: int = 37, ntrials: int = 100000) -> float:
    """
    Plays a game described in q14 ntrials times with a predetermined seed
.
    :param seed: seed for the numpy random number generator.
    :param ntrials: the number of trials to run.
    :return: the probability as described in the written pset.
    """
    np.random.seed(seed)
```

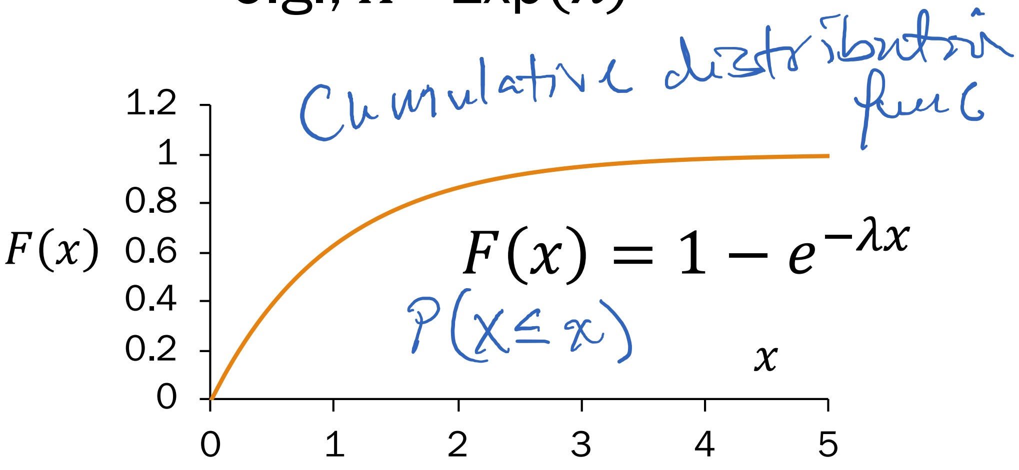
Remember Problem Set 1???

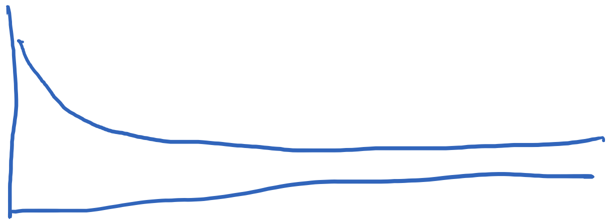
`random.random()` function

`random.random()`
`np.random.random()`
Generate a random float in interval [0.0, 1.0)
$U \sim \text{Uni}(0,1)$



Generate a random number X according to a distribution
e.g., $X \sim \text{Exp}(\lambda)$





Inverse Transform Sampling

Given the ability to generate numbers ~Uni(0,1), how do we generate another number according to a CDF F?

$$X = F^{-1}(U)$$

def F^{-1} the inverse of CDF: $F^{-1}(a) = b \Leftrightarrow F(b) = a$

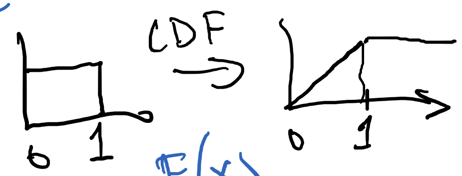
Interpret: If we have a RV $U\sim\text{Uni}(0,1)$, the above RV X (which is a function of U) follows a probability distribution such that $P(X \le x) = F(x)$.

Proof:

$$P(X \le x) = P(F^{-1}(U) \le x)$$

$$= P\big(U \le F(x)\big) \quad (\forall x\colon\ 0 \le F(x) \le 1)$$

$$= F(x) \quad (\text{CDF } P(U \le u) = u \text{ if } 0 \le u \le 1)$$



Inverse Transform Sampling (Continuous)

How do we generate the exponential distribution $X \sim \text{Exp}(\lambda)$?

- CDF: $F(x) = 1 - e^{-\lambda x}$ where $x \geq 0$
- Compute inverse:

$$F^{-1}(u) = -\frac{\log(1-u)}{\lambda}$$

$F(x) = 1 - e^{-\lambda x} = u$

$1 - u = e^{-\lambda x}$

$\log(1-u) = -\lambda x$

$x = \frac{-\log(1-u)}{\lambda}$

- Note if $U \sim \text{Uni}(0,1)$, then $(1-U) \sim \text{Uni}(0,1)$
- Therefore:

$$F^{-1}(U) = -\frac{\log(U)}{\lambda}$$

- Note: Closed-form inverse may not always exist, like with the Normal distribution

Check it out!!! (demo)

Inverse Transform Sampling (Discrete)

$X \sim \text{Poi}(\lambda = 3)$ has CDF $F(X = x)$ as shown:

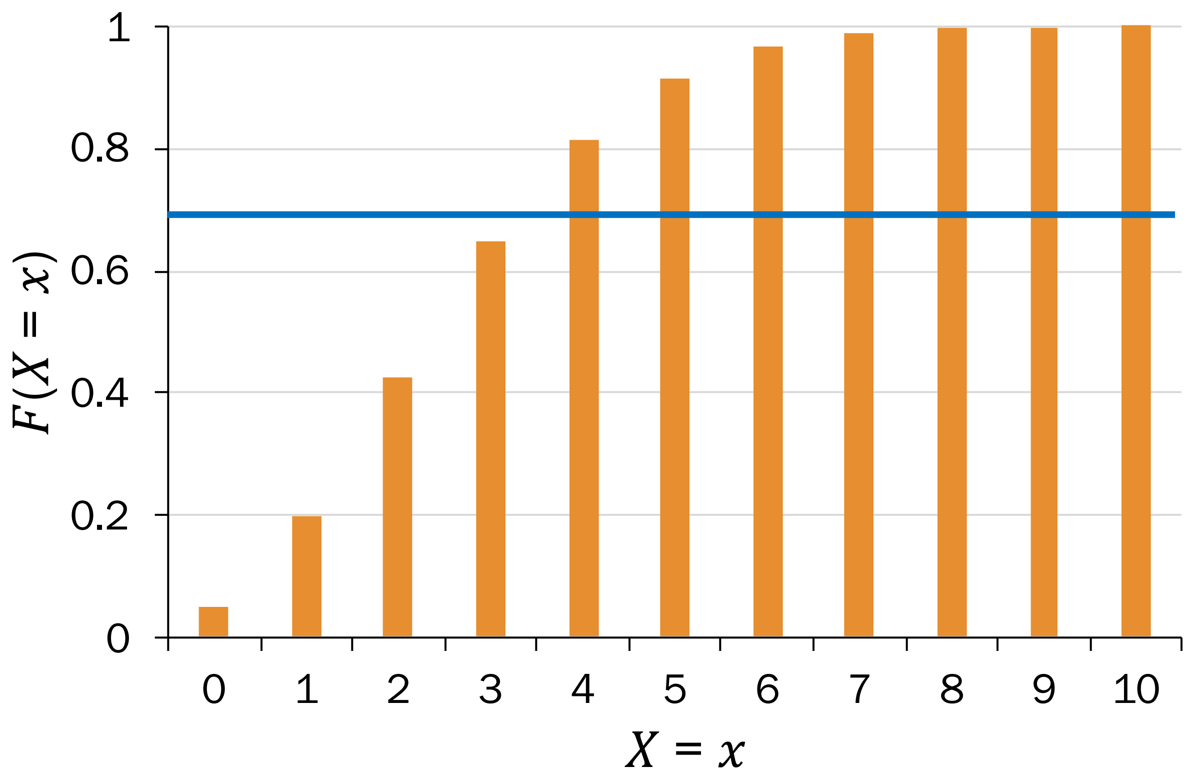
1. Generate $U \sim \text{Uni}(0,1)$

   $u = 0.7$

2. As x increases, determine first $F(x) \geq U$

   $x = 4$

3. Return this value of x



Check it out!!! (demo)

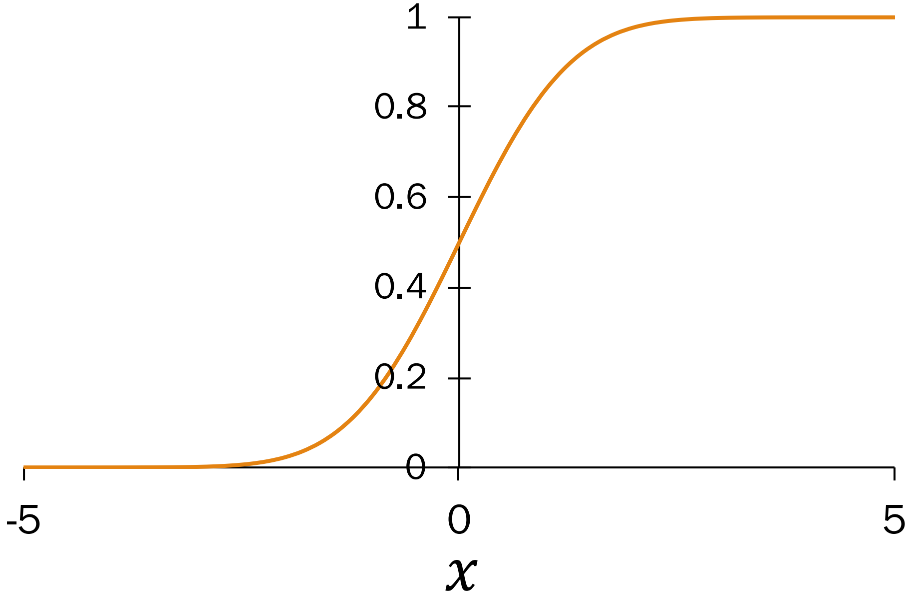
How does a computer sample the Normal?

How does Python generate random values according to a Normal distribution?

```
from scipy import stats
mean = 0
std = 1
for i in range(10):
  sample = stats.norm.rvs(mean, std)
  print(sample)
```

```
-1.5213511002970745
 1.3986457271717916
 2.166196649558274
-0.0961204584265302
-0.650468101242495
-0.661464998510674
-1.1273650614139048
-1.889848256569443
-2.480420257501705
 0.814194996075227
```

CDF of Standard Normal, $\Phi(x)$



Inverse transform sampling

1. Generate a random probability u from $U \sim \text{Unif}(0,1)$.
2. Find x such that $\Phi(x) = u$. In other words, compute $x = \Phi^{-1}(u)$.

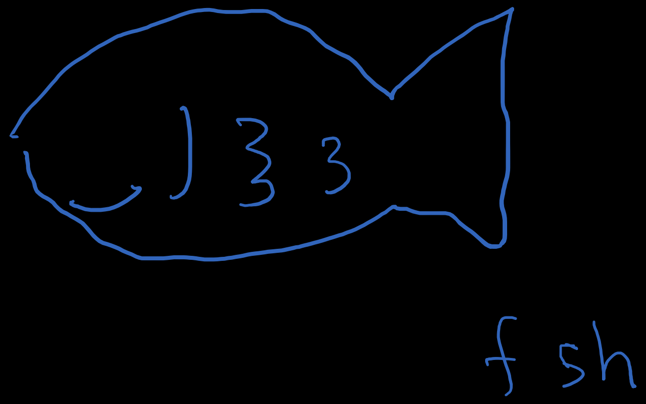
(Since Φ^{-1} has no analytical solution, look up Box-Muller transform for further reading)

https://en.wikipedia.org/wiki/Box%E2%80%93Muller_transform

Another option: Rejection Filtering

Monte Carlo sampling!

Check out Ross 10.2.2 for more information



Interlude for jokes/announcements

Announcements

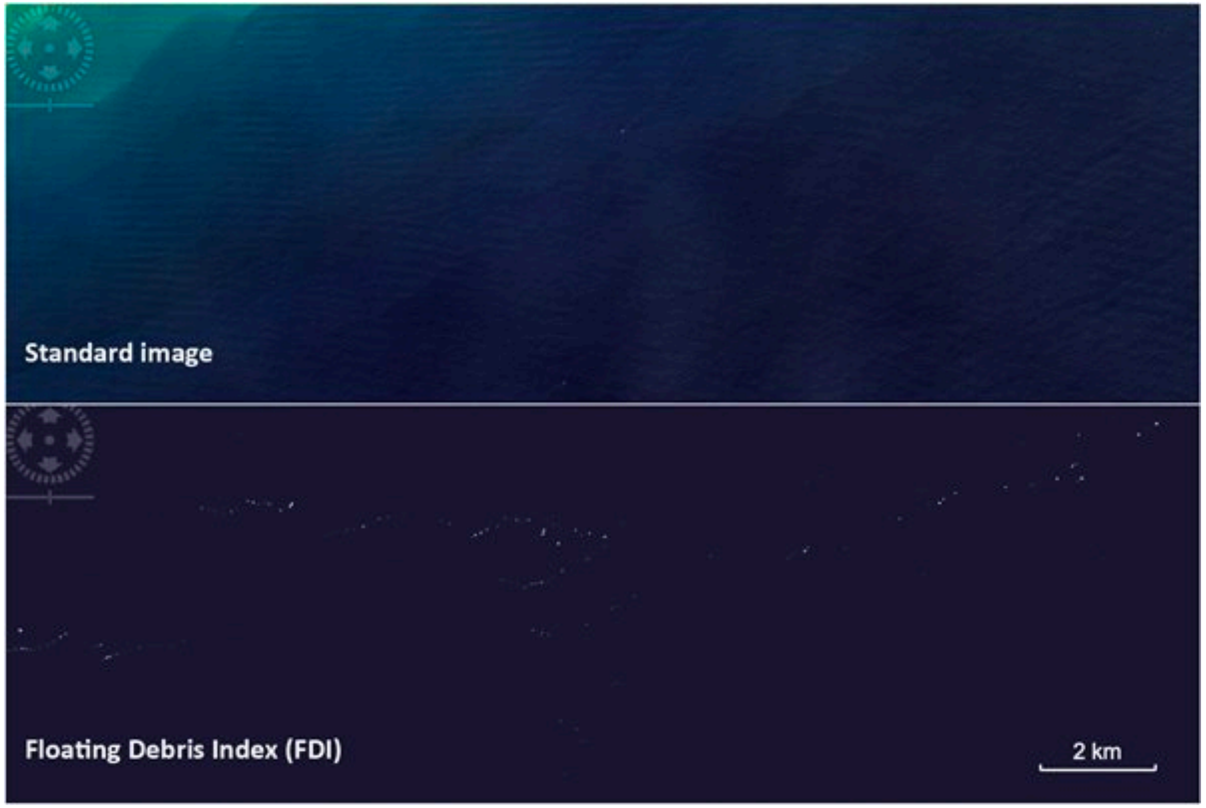
End of Quarter changes, part 2

- Problem Set 6 is now optional (still part of your final weighted course grade, but if you can pass without it, you don't need to turn it in)
- Passing work requirement now 6 of 8 assignments (PS1 to PS6, Quiz 1, Quiz 2)
- ≥ 65%

Interesting probability news

Pioneering technique uses satellites to detect ocean plastic

https://www.circularonline.co.uk/news/pioneering-technique-uses-satellites-to-detect-ocean-plastic/



- Manually, they selected pixels that were suspected to be dominated by plastics using the spectral signature and the FDI, as well as a Normalised Difference Vegetation Index (NDVI).
- Then using an automated approach, floating materials were differentiated using a Naïve Bayes (Bayesian) classification model.
- Across the four study sites, suspected plastics were successfully classified as plastics with an overall accuracy of 86% (Gulf Islands 100%, Accra 87%, Scotland 83% and Da Nang 77%).

Jensen's Inequality

Jensen's inequality

Jensen's inequality:

If $g(x)$ is a convex function, then $E[g(X)] \geq g(E[X])$.

Johan Ludwig William Valdemar Jensen
Danish mathematician
(1859–1925)

Dr. Eggman from *Sonic the Hedgehog*?

Jensen's inequality

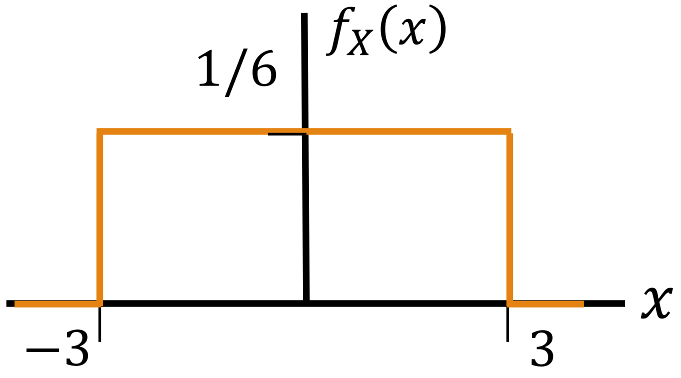
Jensen's inequality:

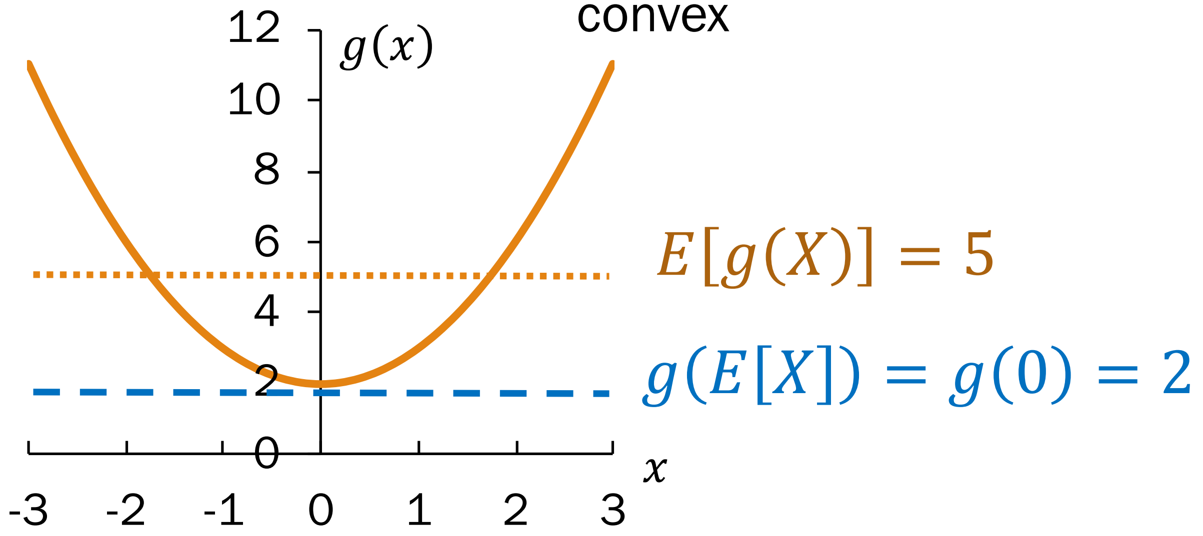
> If $g(x)$ is a convex function , then $E[g(X)] \geq g(E[X])$.

<u>def</u> convex function $g(x)$: if $g''(x) \geq 0$ for all x. (Convex = "bowl")

<u>def</u> concave function $g(x)$: if $-g(x)$ is convex.

Let $X \sim \text{Uni}(-3, 3)$.
Define $g(X) = X^2 + 2$.



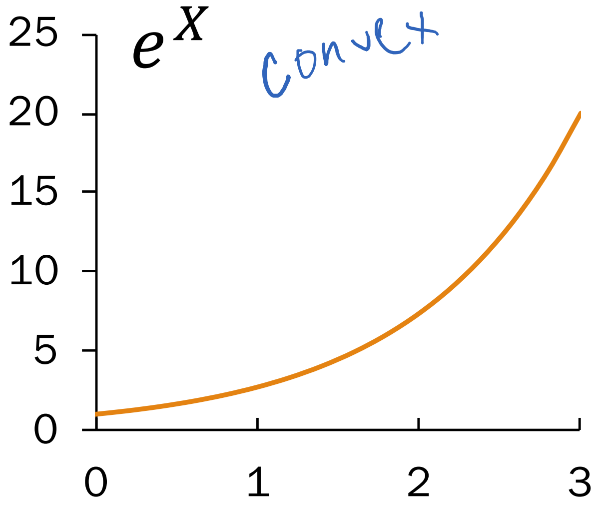


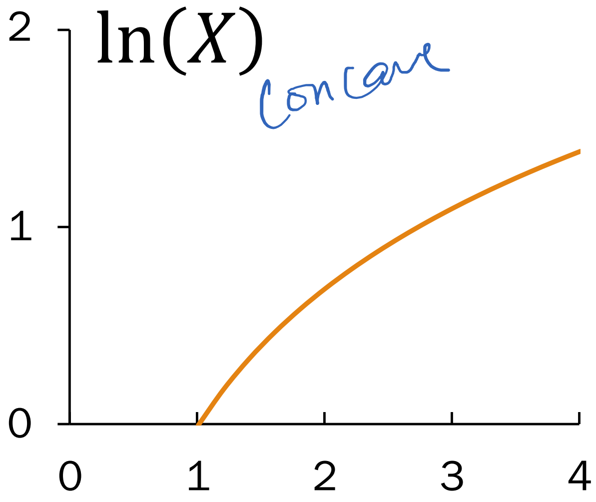
Jensen's quick check

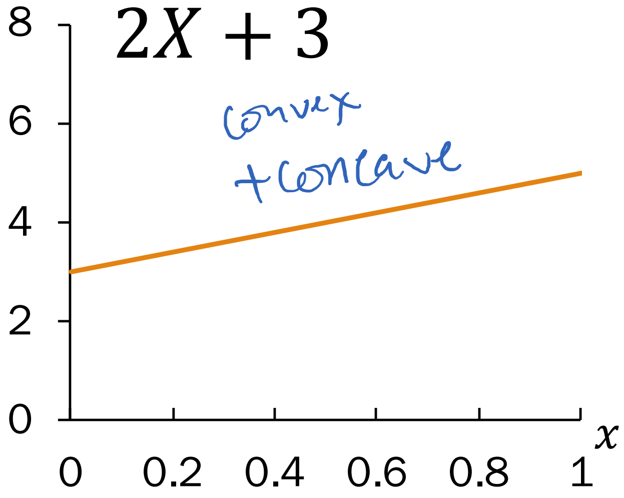
$g(x)$ is convex, $\forall x : g''(x) \geq 0$ ⇒ $E[g(X)] \geq g(E[X])$

Let X~Uniform for the domain of each below graph.

Compare $E[g(X)]$ and $g(E[X])$: (>, <, =)

1.


2.


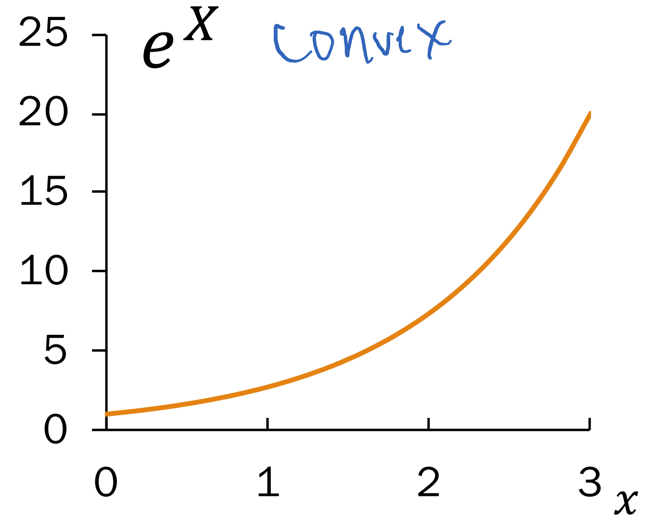
3.


Jensen's quick check

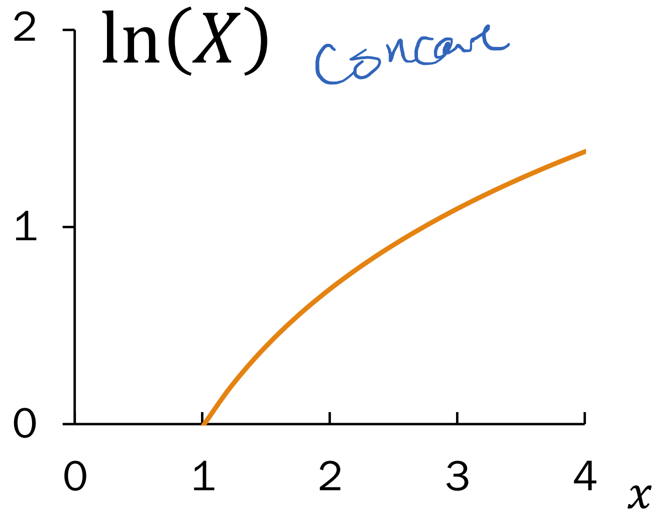
$g(x)$ is convex, $\forall x: \; g''(x) \geq 0$ ⇒ $E[g(X)] \geq g(E[X])$

Let $X \sim$Uniform for the domain of each below graph.

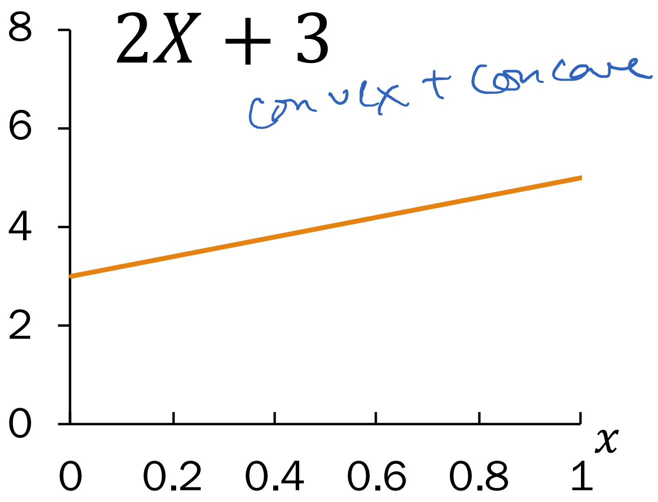
Compare $E[g(X)]$ and $g(E[X])$: (>, <, =)

1. 

$E[e^X] > e^{E[X]}$

2. 

$E[\ln(X)] < \ln(E[X])$

3. 

$E[2X + 3] = 2E[X] + 3$

g is both concave and convex only if it is linear.
$E[g(X)] = g(E[X])$ only if $g(x)$ is a linear function.

Why Jensen's is useful

$g(x)$ is convex, $\forall x: g''(x) \geq 0$ ⇒ $E[g(X)] \geq g(E[X])$

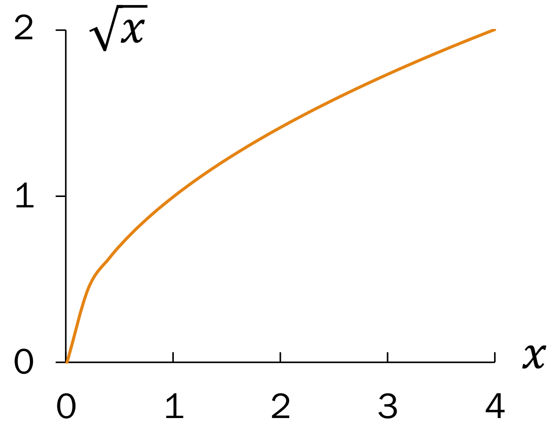
1. Is Standard Error an unbiased estimator? No!

$$E[S^2] = \sigma^2$$ S^2 is an unbiased estimate of σ^2

$$E[S^2/n] = \sigma^2/n$$ Linearity of expectation

$$E\left[\sqrt{S^2/n}\right] < \sqrt{\sigma^2/n}$$ Square root is concavex

$$SE = \sqrt{\frac{S^2}{n}}$$

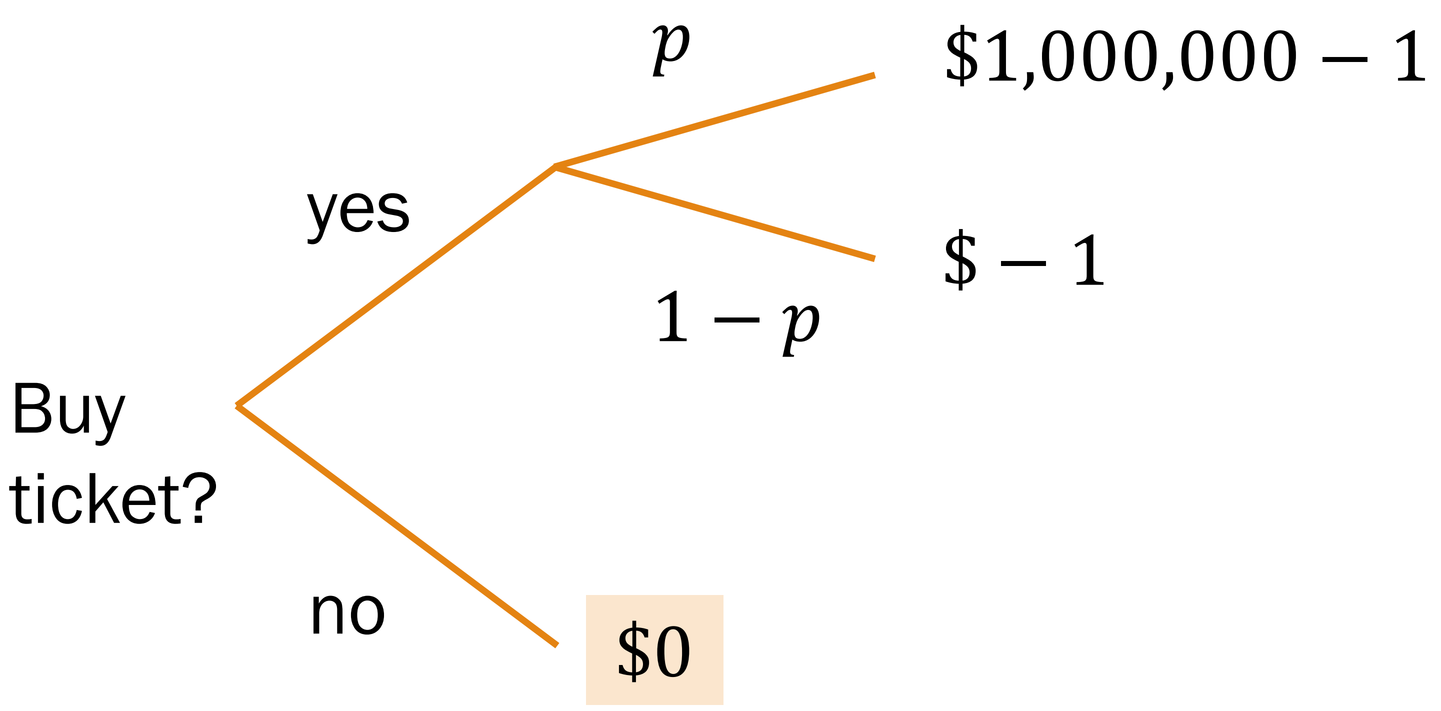


Jensen's Inequality also used in:

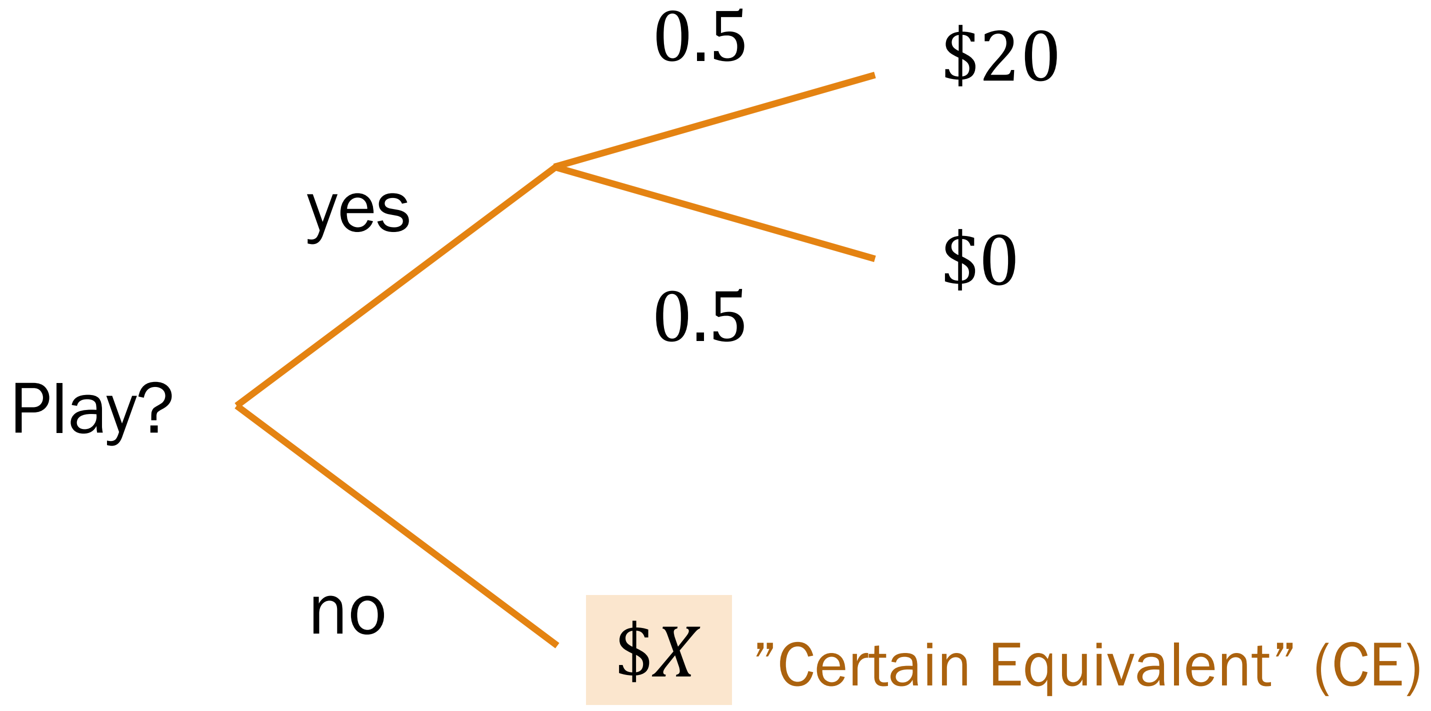
- CS228, KL divergence: What is the best approximate distribution $q(\boldsymbol{x})$ to perform Bayesian inference where the true distribution is actually $p(\boldsymbol{x})$?
- CS229, EM algorithm: How do we iteratively find the the maximum likelihood or MAP estimates without performing gradient ascent?

Utility of Money

Recall the probability tree!



Let's play a game. What choice would you make?



For what value of $\$X$ are you <u>indifferent</u> to playing?

A. $X = 3$
B. $X = 7$
C. $X = 9$
D. $X = 10$

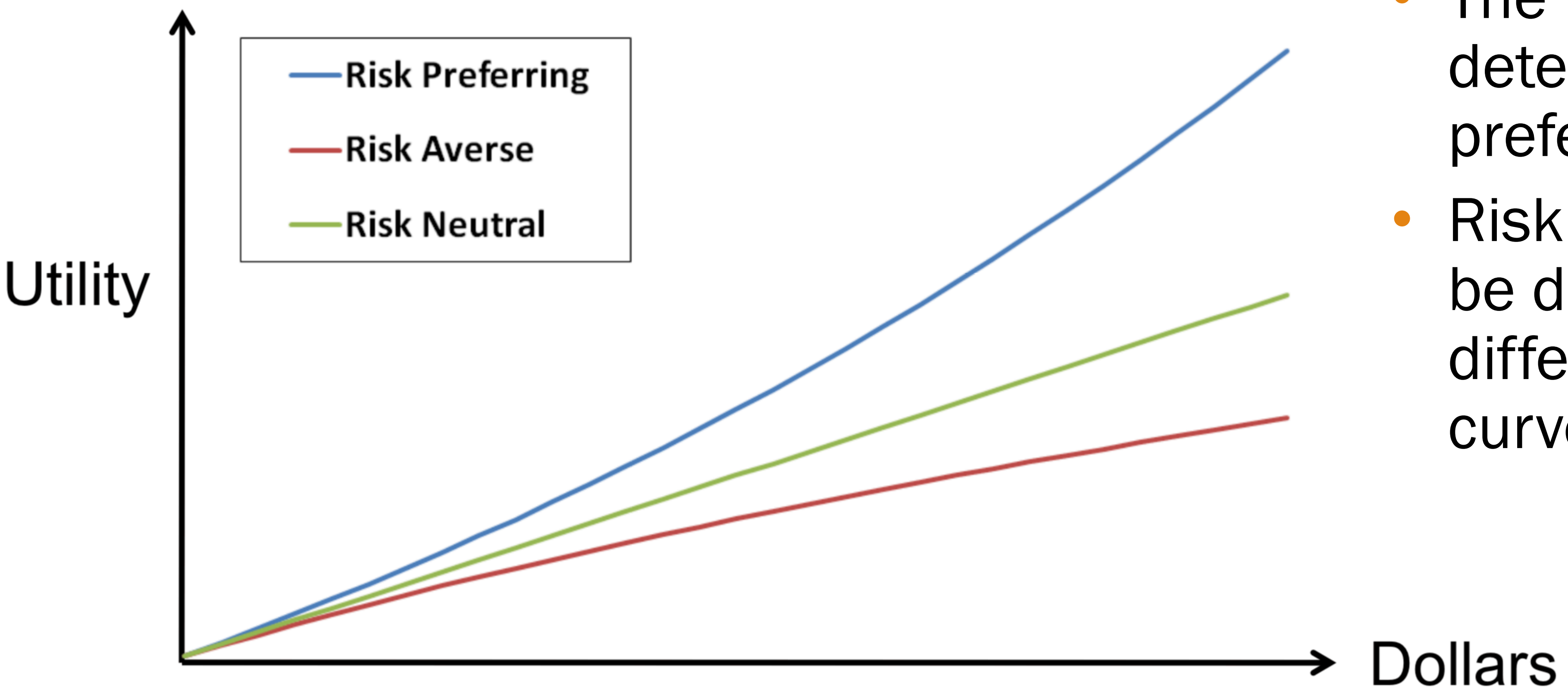
<u>def</u> Certain equivalent: The value of the game to *you* (different for different people)

Utility

<u>def</u> Utility $U(X)$ is the "value" you derive from X

- Can be monetary, but often includes intangibles like quality of life, life expectancy, personal beliefs, etc.

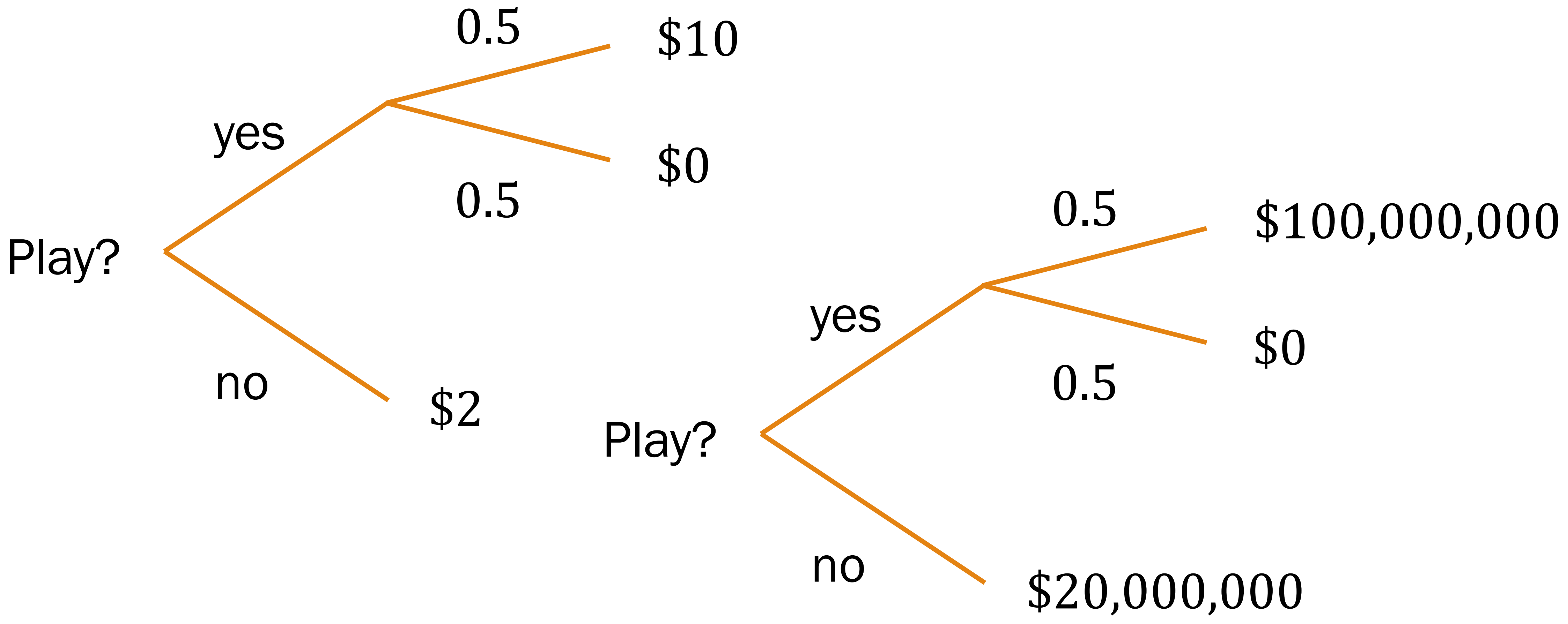
Utility curves



- The utility curve determines your "risk preference."
- Risk preference can be different in different parts of the curve

Non-linearity utility of money

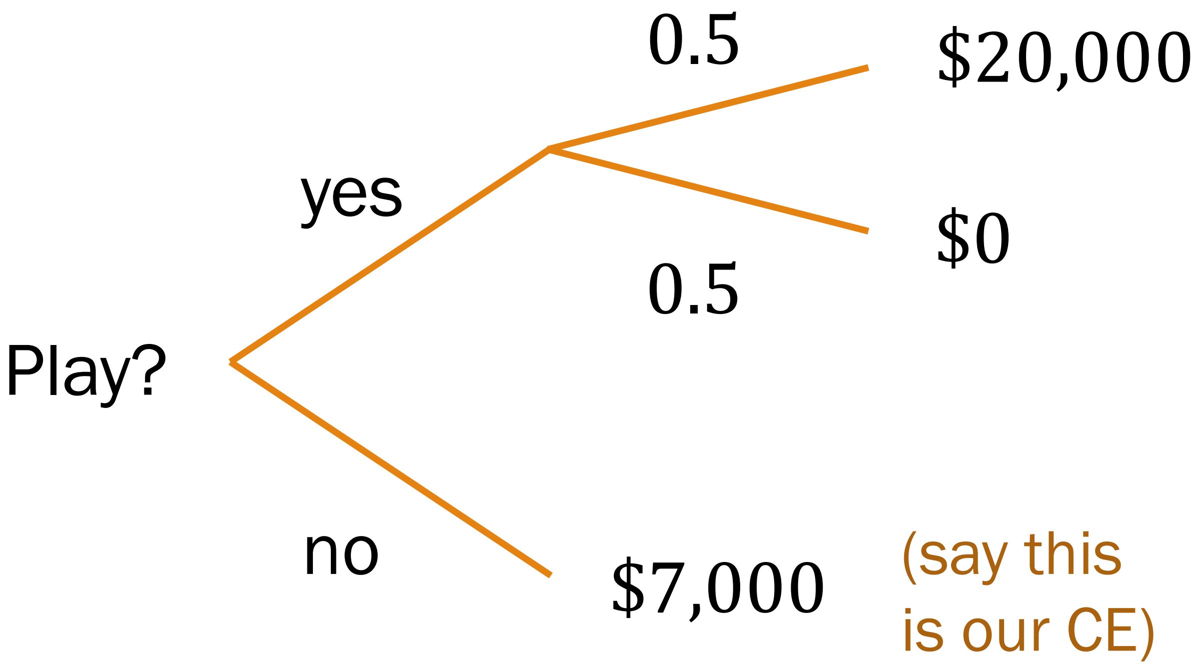
Interestingly, these two choices are different for most people:



Insurance and risk premium

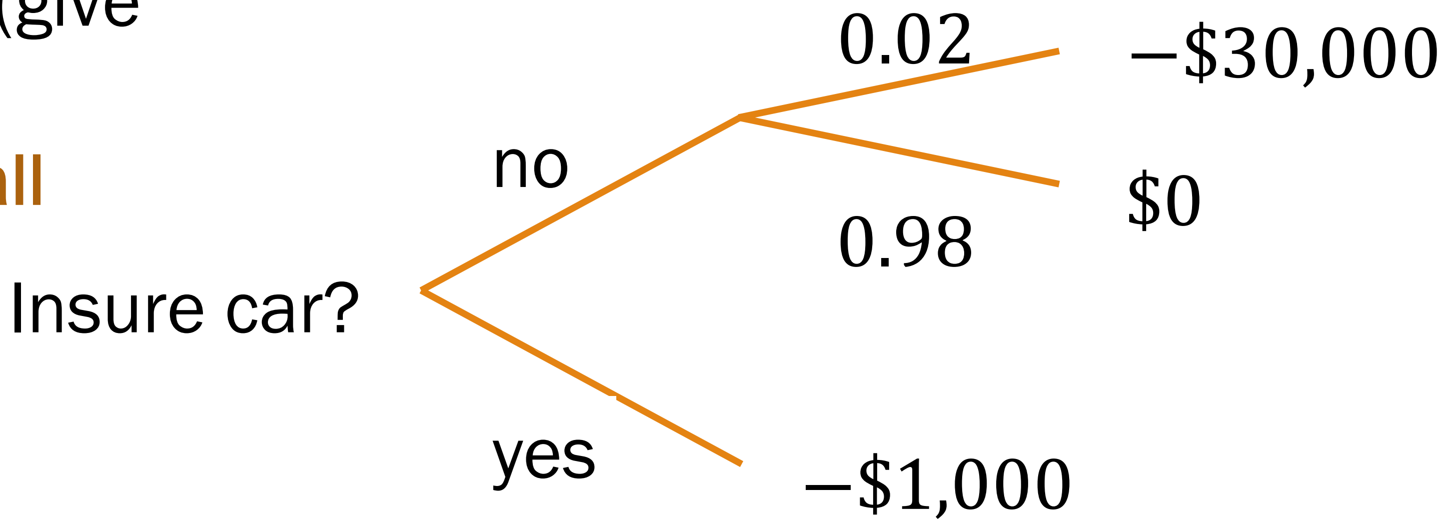
A slightly different game:

- Expected monetary value (EMV) = expected dollar value of game (here, $10,000)



Risk premium = EMV – CE = $3000

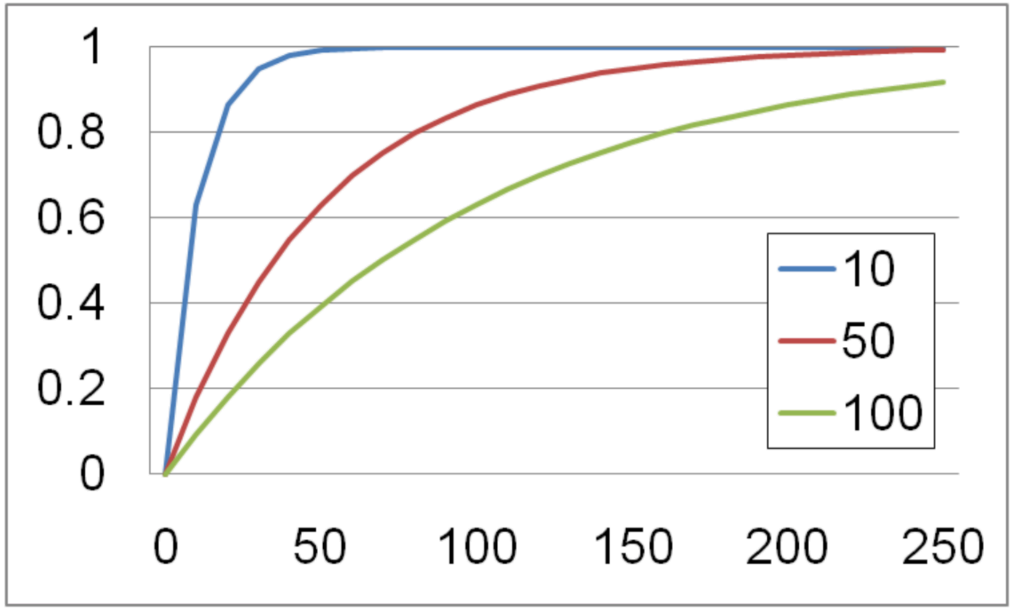
- How much would you pay (give up) to avoid risk?
- **This is what insurance is all about.**



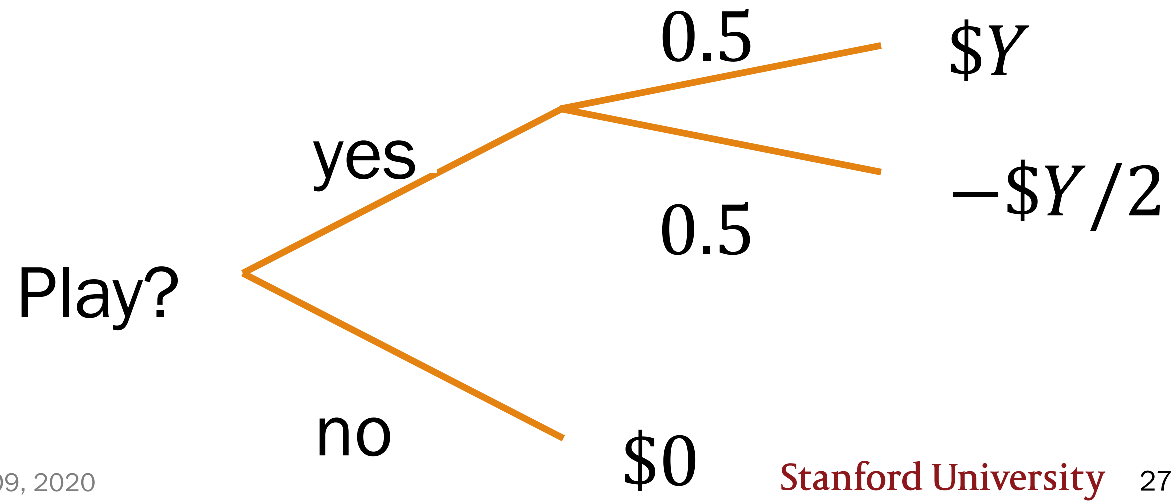
Exponential utility curves

Many people have exponential utility curves:

$$U(x) = 1 - e^{-x/R}$$



- R is your "risk tolerance"
- Larger R = less risk aversion. Makes utility function more "linear"
- $R \approx$ highest value of Y for which you would play:



How rational are you?

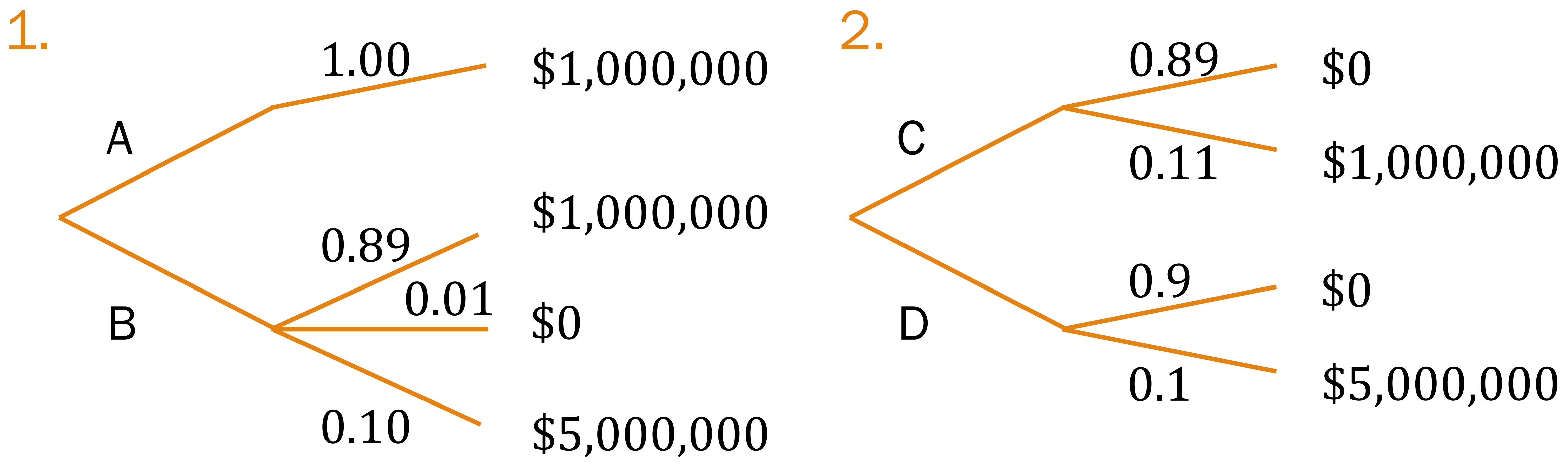
1.

2.

Which option would you choose in each case?

How many of you chose A and D?

How rational are you?



Choice A preferred:

$1.00\ U(1{,}000{,}000) > 0.89\ U(1{,}000{,}000) + 0.01\ U(0) + 0.10\ U(5{,}000{,}000)$

Choice D preferred:

$0.89\ U(0) + 0.11\ U(1{,}000{,}000) < 0.90\ U(0) + 0.10\ U(5{,}000{,}000)$

How rational are you?

Choice D preferred:
$0.11\ U(1{,}000{,}000) < 0.01\ U(0) + 0.10\ U(5{,}000{,}000)$

add $0.89\ U(1{,}000{,}000)$ to both sides

Choice D preferred:
$1.00\ U(1{,}000{,}000) < 0.89\ U(1{,}000{,}000) + 0.01\ U(0) + 0.10\ U(5{,}000{,}000)$

Contradiction???

Choice A preferred:
$1.00\ U(1{,}000{,}000) > 0.89\ U(1{,}000{,}000) + 0.01\ U(0) + 0.10\ U(5{,}000{,}000)$

Choice D preferred:
$0.89\ U(0) + 0.11\ U(1{,}000{,}000) < 0.90\ U(0) + 0.10\ U(5{,}000{,}000)$

subtract $0.89\ U(0)$ from both sides

How rational are you?

Choice D preferred:
$1.00\ U(1{,}000{,}000) <$
$0.89\ U(1{,}000{,}000) +$
$0.01\ U(0) +$
$0.10\ U(5{,}000{,}000)$

add
$0.89\ U(1{,}000{,}000)$
to both sides

Choice D preferred:
$0.11\ U(1{,}000$…

Con…

subtract $0.89\ U(0)$
from both sides

Cho…
1.00 …$000{,}000) >$
$0.89\ U(1{,}000{,}000) + 0.01\ U(0)$
$+0.10\ U(5{,}000{,}000)$

Choice D preferred:
$0.89\ U(0) + 0.11\ U(1{,}000{,}000) <$
$0.90\ U(0) + 0.10\ U(5{,}000{,}000)$

You are inconsistent with utility theory (Allais Paradox)!
⚠️ Human behavior is not always axiomatically consistent ⚠️