A review of basic probability and some loose ends

Different meanings of the = symbol

Definition: $E(X) \doteq \sum_{i=1}^{\infty} s_i P(X = s_i)$

Equality: $\mathrm{a}^2 + b^2 = (a+b)(a-b)$

$$E(X+Y) = E(X) + E(Y)$$

Assignment: $x \coloneqq x+1$

Equality test: $x == y$: 1 if x=y, 0 otherwise.

"=" as assignment

Exchanging the values of A and B

Using a third variable:

```
A:=B
C:=A
B:=C
```

In Place:

```
A:=A+B
B:=A-B
A:=A-B
```

Order is important!

`A:=A+1` ***makes sense!***

"=" as Mathematical equality

$$(X+Y)^2 = X^2 + 2XY + Y^2$$

Holds for all values of X and Y : Universal truth

Holds if X and Y are random variables

$Y = 1$: constrains Y to a single value

$Y = X^2$: constrains allowed (X,Y)

$X = 3$: constrains X to a single value

But $1 \neq 3^2$ *CONTRADICTION*

All constraints hold simultanously.
Order of constraints does not matter !!
X=3 means the same thing as 3=X !!

A=A+1 Contradiction (no satisfying A)

"=" in definitions

$$E(X) \doteq \sum_{i=1}^{\infty} s_i \Pr(X(\omega) = s_i)$$

$$\mu \doteq E(X)$$

$$\mathrm{var}(X) \doteq E\left[(X - \mu)^2\right]$$

Substituting the definition of E(X):

$$\mathrm{var}(X) = \sum_{i=1}^{\infty} (s_i - \mu)^2 \Pr\left(X(\omega) = s_i\right)$$

Definitions can come in any order, but the convention is that complex definitions come after the simpler definitions on which they are based.

Independent random variables

$$\forall x, y \quad P(X = x \wedge Y = y) = P(X = x)P(Y = y)$$

Consider two values for Y : $y_1 \neq y_2$

Then: $\forall x \qquad \dfrac{P(X = x \wedge Y = y_1)}{P(X = x \wedge Y = y_2)} = \dfrac{P(Y = y_1)}{P(Y = y_2)}$

Similarly for any two values of X

Conditions for independence:

* Any two rows in the contingency matrix are multiples of each others

* Any two columns in the contingency table are multiples of each other.

* The contingency table has rank one.

$$Cov(X,Y) \doteq E\left[\left(X-E[X]\right)\left(Y-E[Y]\right)\right] = E[XY]-E[X]E[Y]$$

$$E[XY] = E[X]E[Y] + Cov(X,Y)$$

$$Corr(X,Y) \doteq \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}} = \frac{Cov(X,Y)}{\sigma(X)\sigma(Y)} \Rightarrow \quad \begin{array}{c} Corr(aX,bY) = Corr(X,Y) \\ \text{if } a,b>0 \end{array}$$

Unlike the Covariance, the Correlation Coefficient is unit-less,
Changing the units, or multiplying each random variable by some constant,
does not change the correlation coefficient.
The correlation Coefficient is always in the range $[-1,+1]$

$$X,Y \textit{ independent} \quad \Rightarrow \quad Cov(X,Y) = Corr(X,Y) = 0$$

$$Cov(X,Y) = Corr(X,Y) = 0 \quad \not\Rightarrow \quad X,Y \textit{ independent}$$

Examples

Correlated Variables

	X=1	X=2	X=3	X=4
Y=1	1/4	1/4	0	0
Y=2	0	0	0	0
Y=3	0	0	1/4	1/4

$$\mu(X) = 2.5,\ \mu(Y) = 2$$

$$\text{cov}(X,Y) = \frac{1}{4}(-1.5 * -1) + \frac{1}{4}(-.5 * -1) + \frac{1}{4}(.5 * 1) + \frac{1}{4}(1.5 * 1) = 1$$

Anti Correlated Variables

	X=1	X=2	X=3	X=4
Y=1	0	0	0	1/4
Y=2	0	1/4	1/4	0
Y=3	1/4	0	0	0

$$\mu(X) = 2.5,\ \mu(Y) = 2$$

$$\text{cov}(X,Y) = \frac{1}{4}(-1.5 * 1) + \frac{1}{4}(-.5 * 0) + \frac{1}{4}(.5 * 9) + \frac{1}{4}(1.5 * -1) = -\frac{3}{4}$$

Uncorrelated and independent

	X=1	X=2	X=3	X=4
Y=1	1/4	0	0	1/4
Y=2	0	0	0	0
Y=3	1/4	0	0	1/4

$$\mu(X) = 2.5,\ \mu(Y) = 2$$

$$\operatorname{cov}(X,Y) = \frac{1}{4}(-1.5*1) + \frac{1}{4}(-1.5*-1) + \frac{1}{4}(1.5*1) + \frac{1}{4}(1.5*-1) = 0$$

P(X=1)=P(X=4)=1/2, P(Y=1)=P(Y=3)=1/2

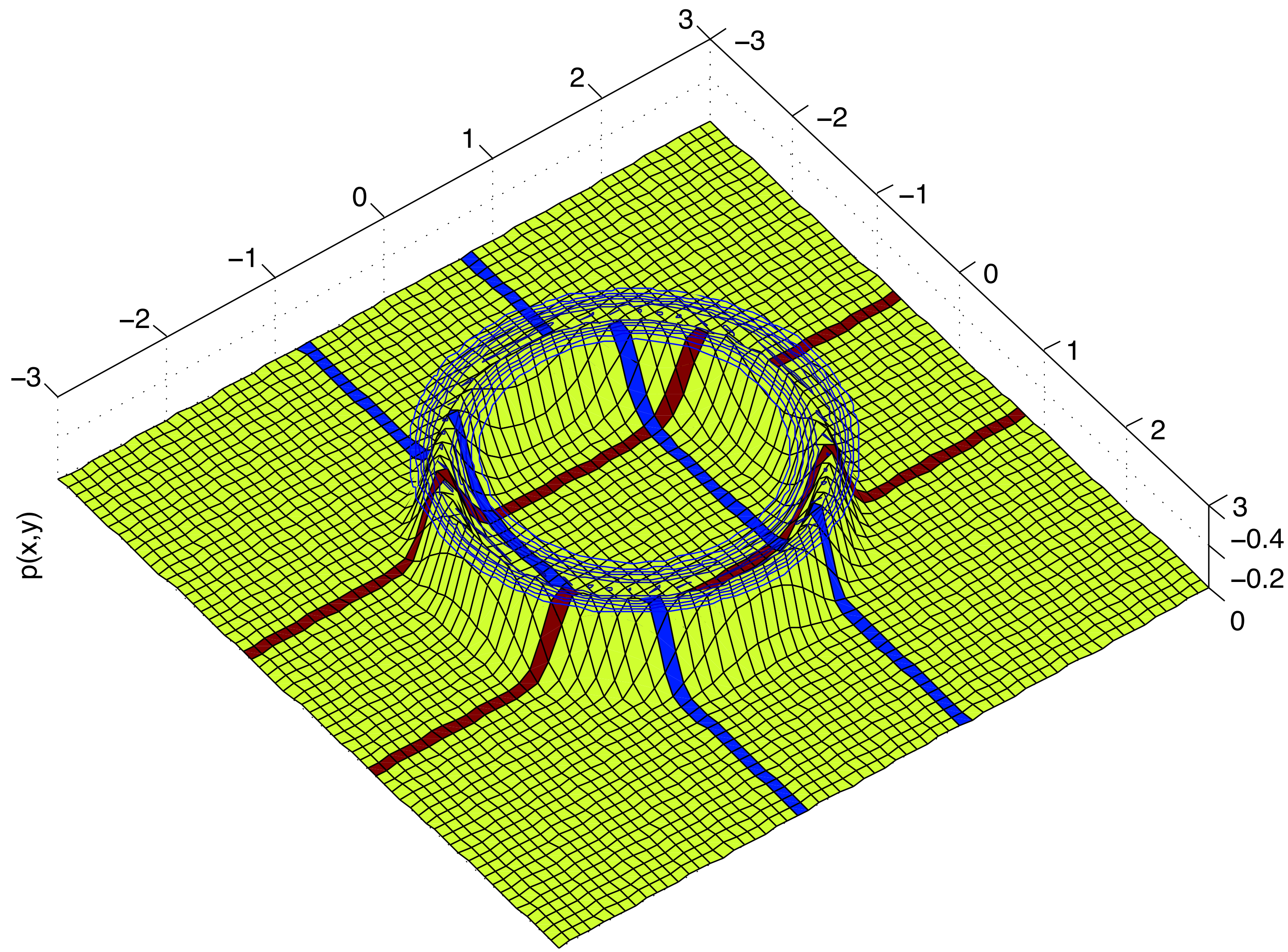
X and Y are independent because all of the joint probabilities are either 0 or 1/4

Uncorrelated but dependent

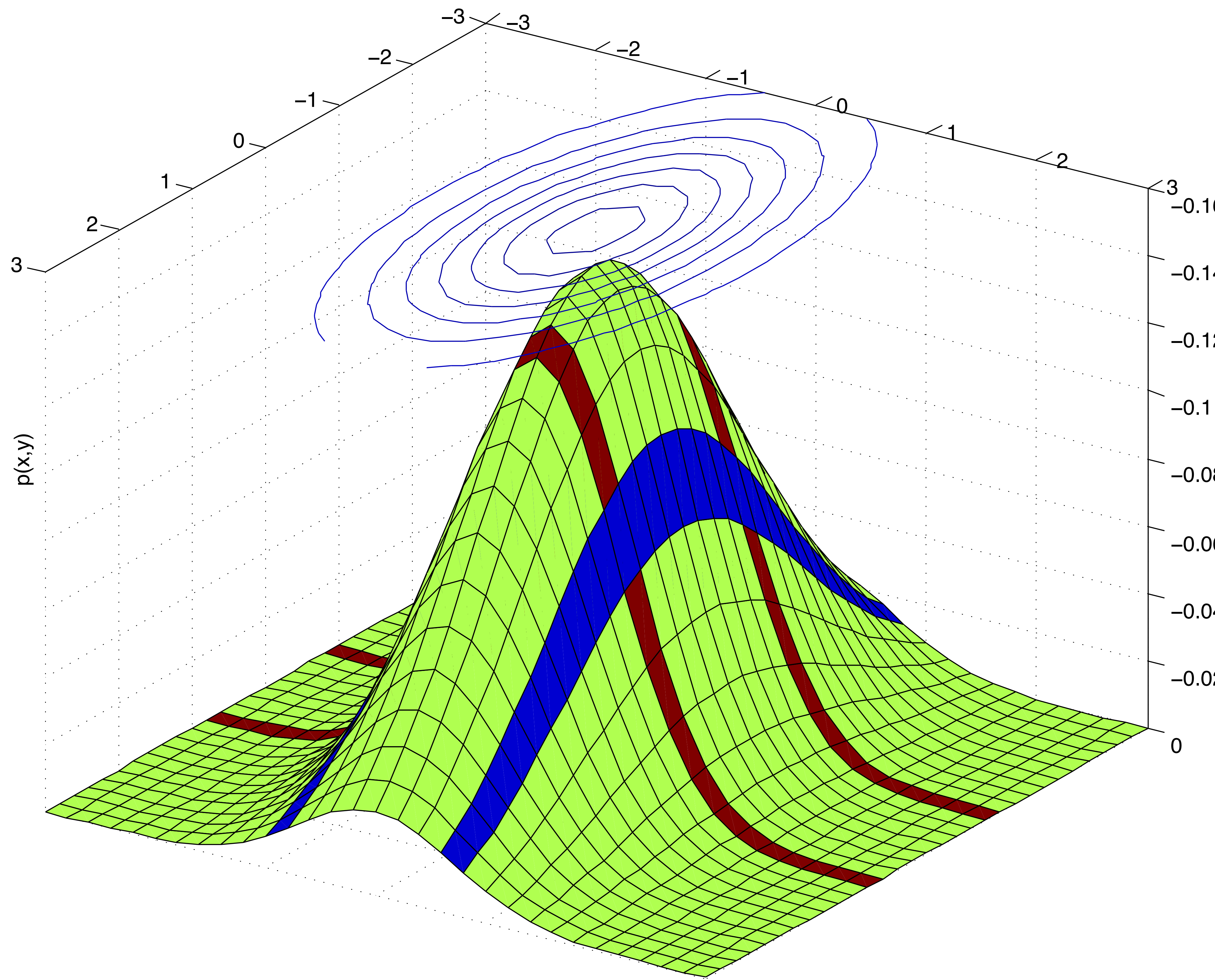
	X=1	X=2	X=3	X=4
Y=1	1/8	0	0	1/8
Y=2	0	1/4	1/4	0
Y=3	1/8	0	0	1/8

1. Cov(X,Y)=0
2. X and Y are independent

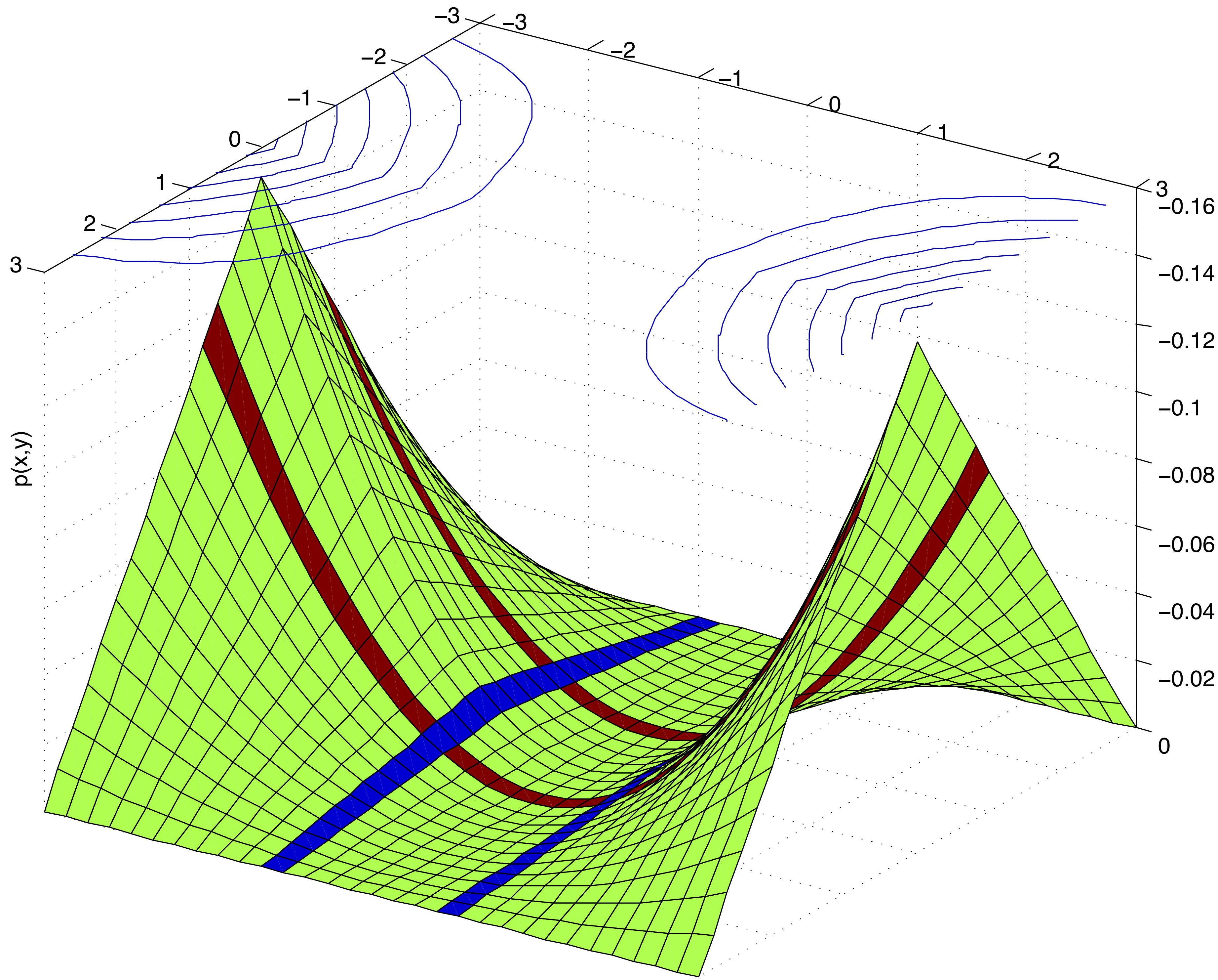
Example 5, Circle



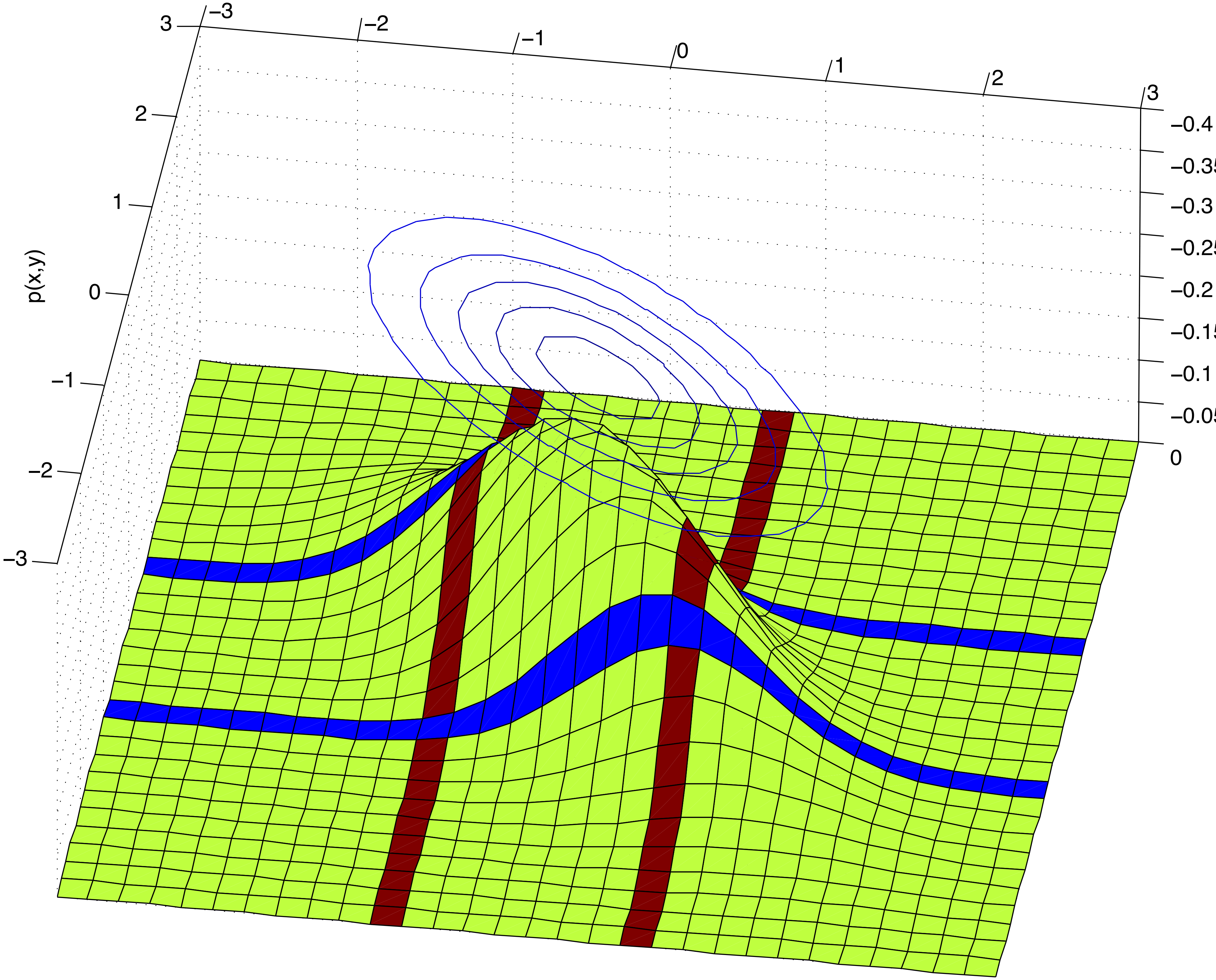
Example 2 independent RVs



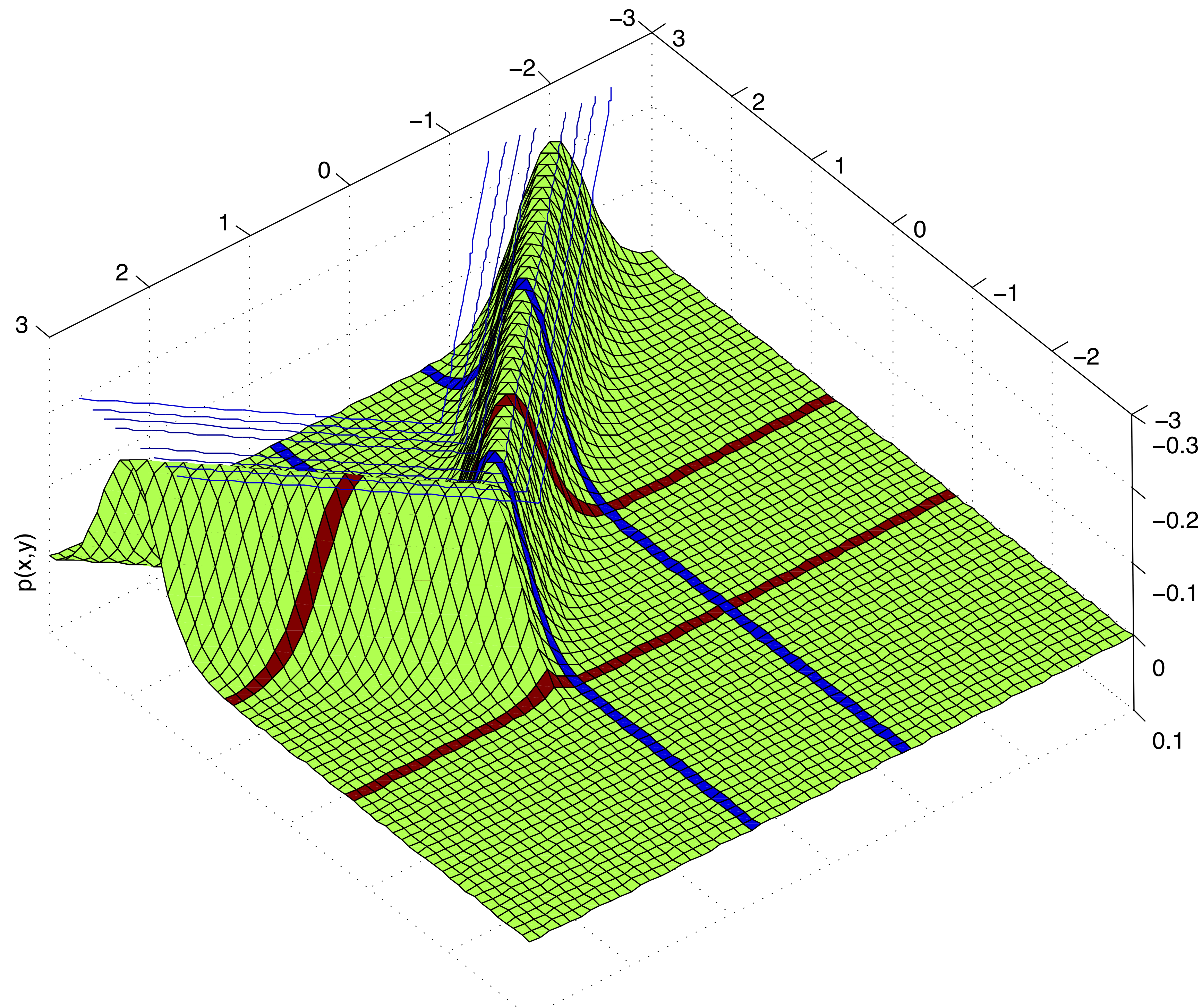
Example 1, independent RVs



Example 3, Dependent RVs



Example 4, functional dependence



Important Statistical Equalities (these are all mathematical equalities)

for any random variables X, Y

and any constants a, b

$- -$

$E(aX) = aE(X); \quad E(X+b) = E(X) + b$

$E(X+Y) = E(X) + E(Y)$

If X, Y are independent: $E(XY) = E(X)E(Y)$

$- -$

$\mathrm{var}(aX) = a^2\,\mathrm{var}(X); \quad \mathrm{var}(X+b) = \mathrm{var}(X)$

If X, Y are independent: $\mathrm{var}(X+Y) = \mathrm{var}(X) + \mathrm{var}(Y)$

Linearity of variance

If X and Y are independent random variables, then $\mathrm{var}(X+Y) = \mathrm{var}(X) + \mathrm{var}(Y)$. More generally, if $X_1, \ldots, X_n$ are independent, then

$$\mathrm{var}(X_1 + \cdots + X_n) \; = \; \mathrm{var}(X_1) + \cdots + \mathrm{var}(X_n).$$

In contrast, linearity of expectation ($\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y)$) holds even if the random variables are *not* independent.

Suppose X and Y are independent random variables. X has mean of 1 and variance of 2, Y has mean of 3 and variance of 2, then for the r.v. $Z = 3X + 4Y + 2$, the mean is ______ and the variance is ______ .

Linearity of variance

If X and Y are independent random variables, then $\mathrm{var}(X+Y) = \mathrm{var}(X) + \mathrm{var}(Y)$. More generally, if $X_1, \ldots, X_n$ are independent, then

$$\mathrm{var}(X_1 + \cdots + X_n) = \mathrm{var}(X_1) + \cdots + \mathrm{var}(X_n).$$

In contrast, linearity of expectation ($\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y)$) holds even if the random variables are *not* independent.

Suppose X and Y are independent random variables. X has mean of 1 and variance of 2, Y has mean of 3 and variance of 2, then for the r.v. $Z = 3X + 4Y + 2$, the mean is ______ and the variance is ______ .

E(3X+4Y+2)=E(3X)+E(4Y)+E(2)=3E(X)+4E(Y)+2=3*1+4*3+2=17

var(3X+4Y+2)=var(3X+4Y)=var(3X)+var(4Y)=9var(X)+16var(Y)
= 9*2+16*2=50

properties of the mean:

X is a random variable getting the values $\{s_1, s_2, \ldots\}$

$$E(aX+b) = \sum_{i=1}^{\infty}(as_i + b)\Pr(X = s_i) = a\sum_{i=1}^{\infty} s_i \Pr(X = s_i) + b\sum_{i=1}^{\infty} \Pr(X = s_i) = aE(X) + b$$

$$X \in \{s_1, s_2, \ldots\}, Y \in \{t_1, t_2, \ldots\}$$

$$E(X+Y) = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty}(s_i + t_j)\Pr(X = s_i, Y = t_j)$$

$$= \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} s_i \Pr(X = s_i, Y = t_j) + \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} t_j \Pr(X = s_i, Y = t_j)$$

$$= \sum_{i=1}^{\infty} s_i \Pr(X = s_i) + \sum_{j=1}^{\infty} t_j \Pr(Y = t_j) = E(X) + E(Y)$$

Linearity of Expectation

Using Linearity of expectation 1

A dice is tossed 3 times, what is the expected value of the sum?

Let $X1, X2, X3$ be the outcomes of the three tosses
$E(X1+X2+X3)=E(X1)+E(X2)+E(X3) = 3E(X1) = 3.5 \times 3 = 10.5$

Using Linearity of expectation 2

Linearity of expectation holds whether or not the RVs are independent

Suppose we flip a fair coin 100 times. What is the expected number of times that we get the pattern HHH? How about for the pattern HTH?

Clearly there are strong dependencies between neighboring patterns:

HTHTHHHHTHTTH

computing the probability of i occurances of the pattern is complicated....

However, let the RV X_i ***be 1 if the pattern appears in positions*** $(i,i+1,i+2)$, ***0 otherwise.***

Clearly the sum $X_1+X_2+\dots+X_{98}$ ***is the number of times that the pattern occured.***

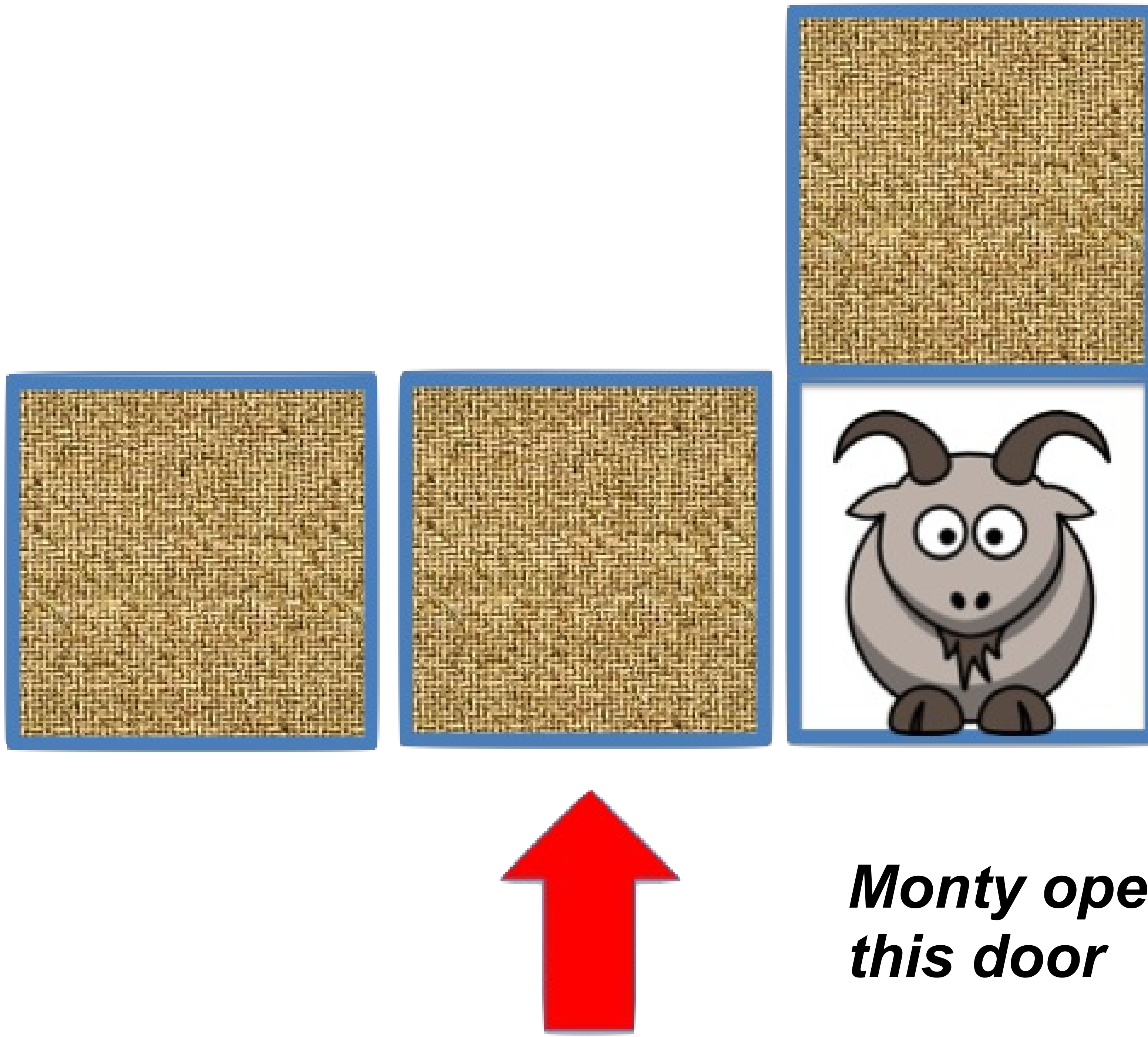
By Linearity of expectation we have
$E(X_1+X_2+\dots+X_{98})=E(X_1)+\dots+E(X_{98})= 98E(X_1)$

But each pattern of length three is equiprobable so $E(X_1)=1/8$

Therefor $E(X_1+X_2+\dots+X_{98})=98/8$ ***for ANY PATTERN of length 3***

The Monty Hall Problem


I am betting
on this door



Monty opens this door

I am betting on this door

I am allowed to switch, should I?

Argument that it does not matter:

The chance that the treasure is behind each of the doors 50%.

As the probabilities are equal, it does not matter whether we switch or not.

Argument for choosing one of the two unopen doors at random.

Before I had to choose between 3 doors - my probability of sucsses was 1/3

Now I am choosing between between two doors, my probability of success is 1/2 So random is better than staying on the same door.

Argument for Switching.

The probability that the treasure is behind the door I chose did not change. Therefor the probability that switching will put me on the treasure must be 2/3:

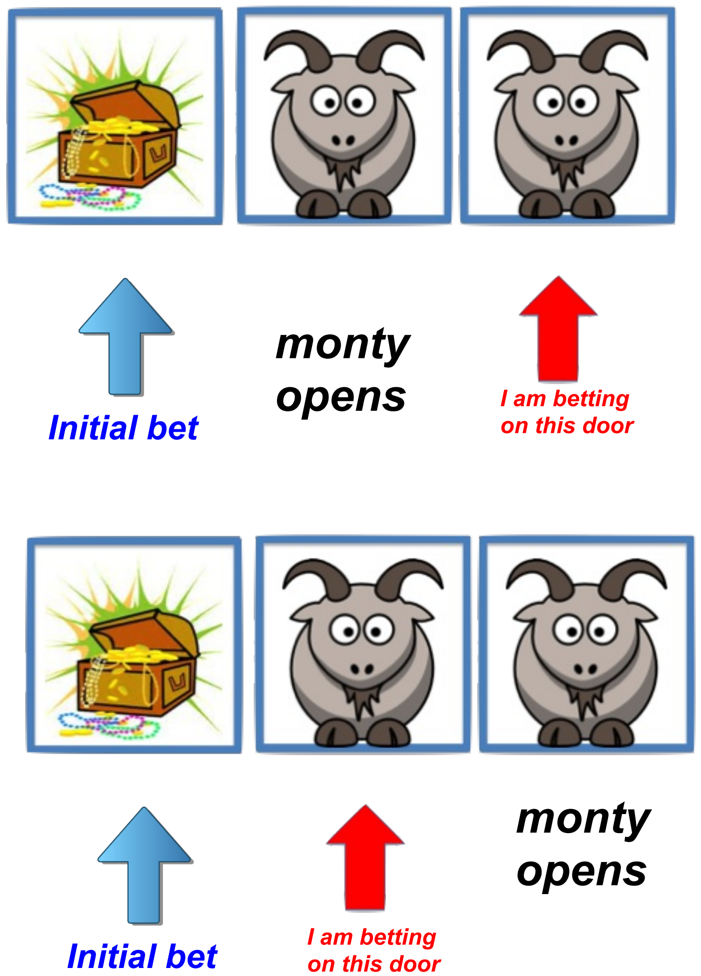
$$1/2*1/3+1/2*2/3 = 1/2$$

Arguments against switching:

I know already that one of the other doors has a goat behind it. So getting the information does not tell me anything new.

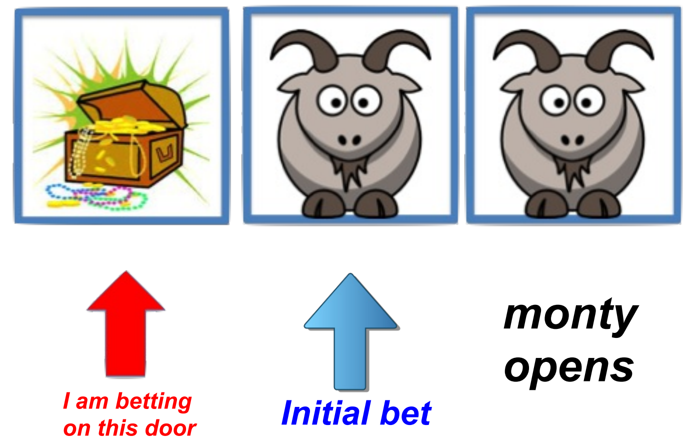
Analysis for always switching

prob 1/3



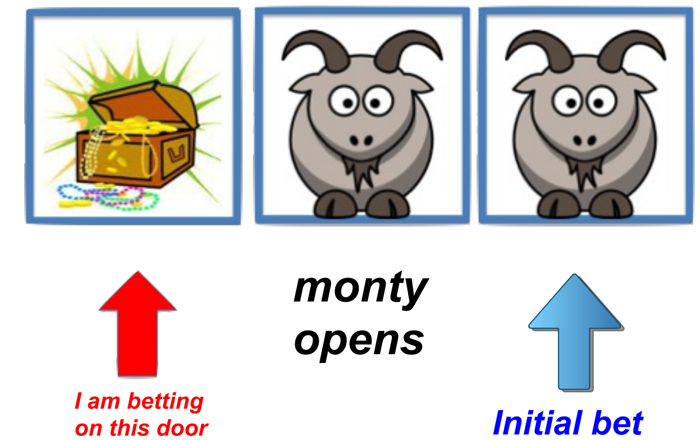
I lose

prob 1/3



I win!

prob 1/3



I win!

Hidden Assumption: ***monty always opens a door to reveal a goat.***

In fact, he might have his own goals:

If Monty wants us to lose: ***open door only when we choose the treasure door.***

If Monty wants us to win: ***open door only when we choose a goat door.***

For us the only SAFE thing to do is not to switch.

This is called the "Min-Max" strategy.

Min-Max is the strategy the guarantees us the best outcome in the worst case.

More on that - game theory.

Expected value over countably infinite sets

$S =$ a countably infinite subset of R

$S = \{s_1, s_2, \ldots\}$

$X =$ a random variable which gets values in S

$$E(X) \doteq \sum_{i=1}^{\infty} s_i \Pr(X(\omega) = s_i)$$

Consider the distribution

distribution is over the natural numbers = positive integers

$$P(X=i)=\frac{1}{zi^3}; \quad z=\sum_{i=1}^{\infty}\frac{1}{i^3}<\infty$$

$$E(X)=\sum_{i=1}^{\infty}\frac{i}{zi^3}=\sum_{i=1}^{\infty}\frac{1}{zi^2}<\infty$$

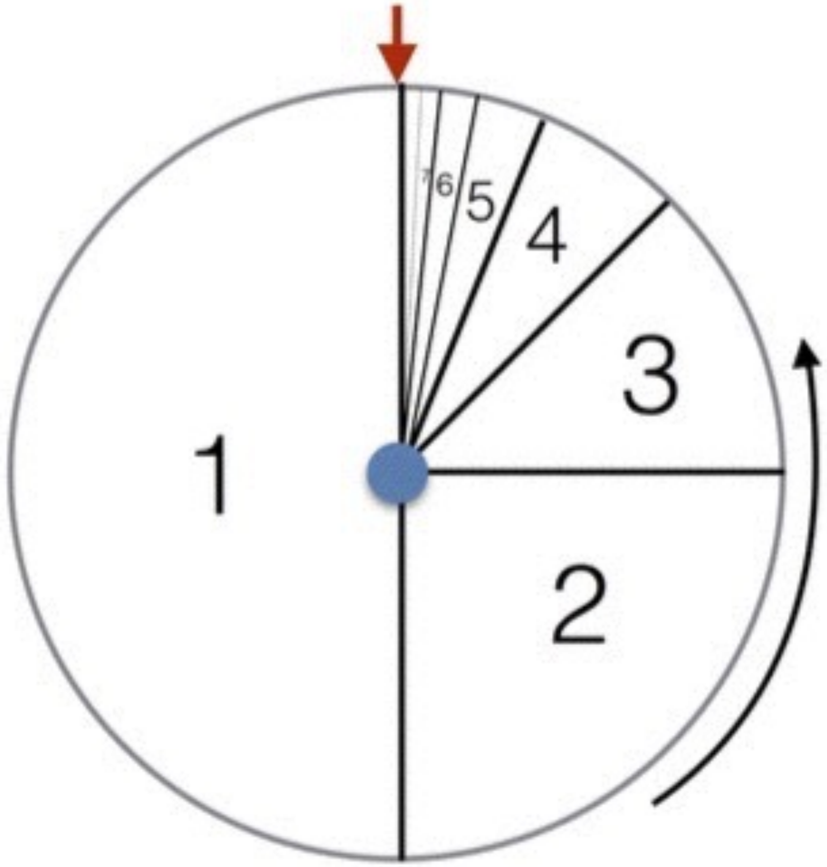
Expectation is finite

Consider next the distribution

$$P(X=i)=\frac{1}{zi^2}; \quad z=\sum_{i=1}^{\infty}\frac{1}{i^2}<\infty$$

$$E(X)=\sum_{i=1}^{\infty}\frac{i}{zi^2}=\sum_{i=1}^{\infty}\frac{1}{zi}=\infty$$

Distribution is well defined but Expectation is infinite



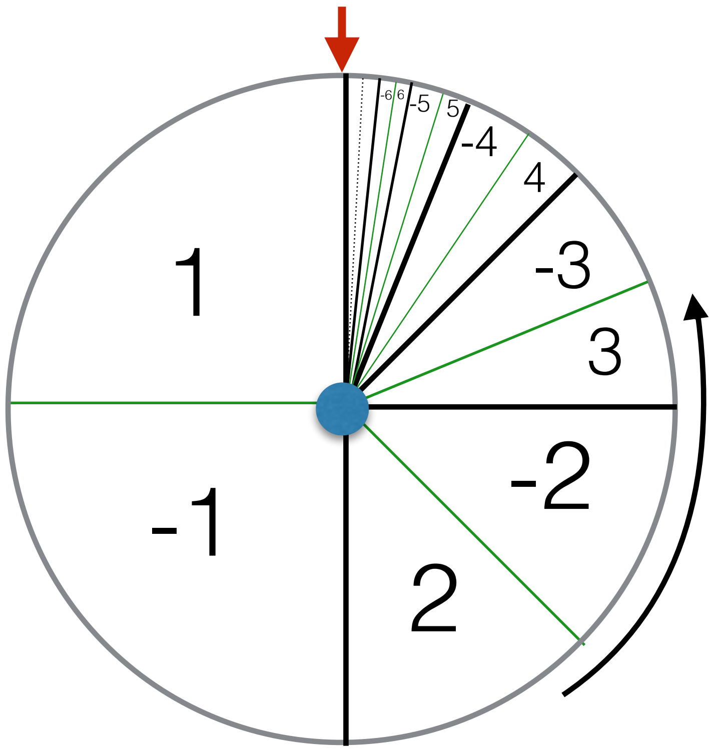
(b) A wheel with
Infinitely many outcomes

$$P(X=i)=\frac{6}{\pi^2 i^2};$$

$$\sum_{i=1}^{\infty} P(X=i)=1$$

$$\sum_{i=1}^{\infty} iP(X=i)=\infty$$

Participation in this game is worth any price
(on the long term)



A wheel with
Infinitely many outcomes
both positive and negative

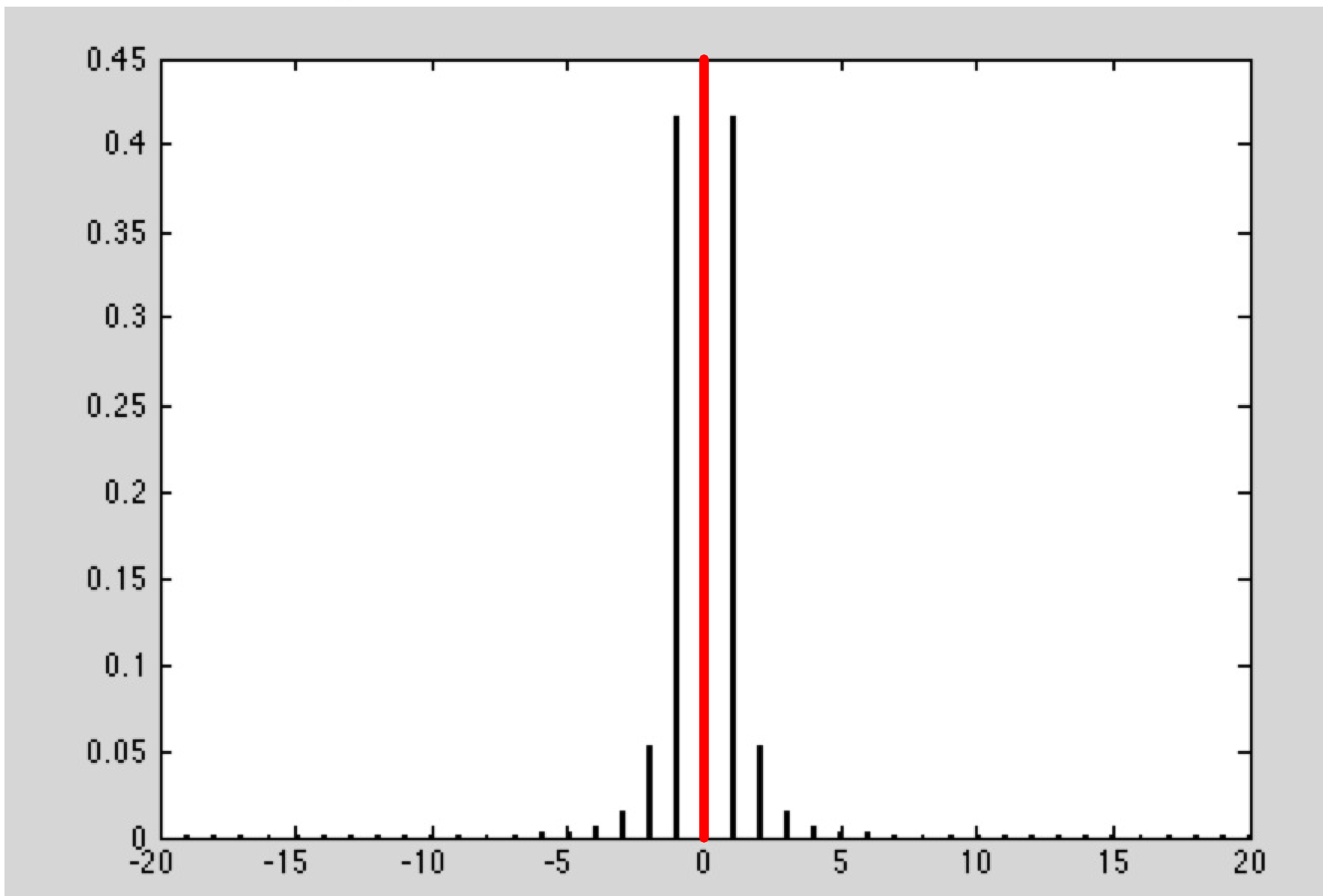
$$P(X=i)=\begin{cases} \dfrac{1}{i^{1.5}} & if \quad i \neq 0 \\ 0 & if \quad i = 0 \end{cases}$$

$$\sum_{i=-\infty}^{\infty} P(X=i) = 1$$

$$\sum_{i=\infty}^{\infty} iP(X=i) \ \ is\ undefined$$

Expectation over pos and neg integers: the good case

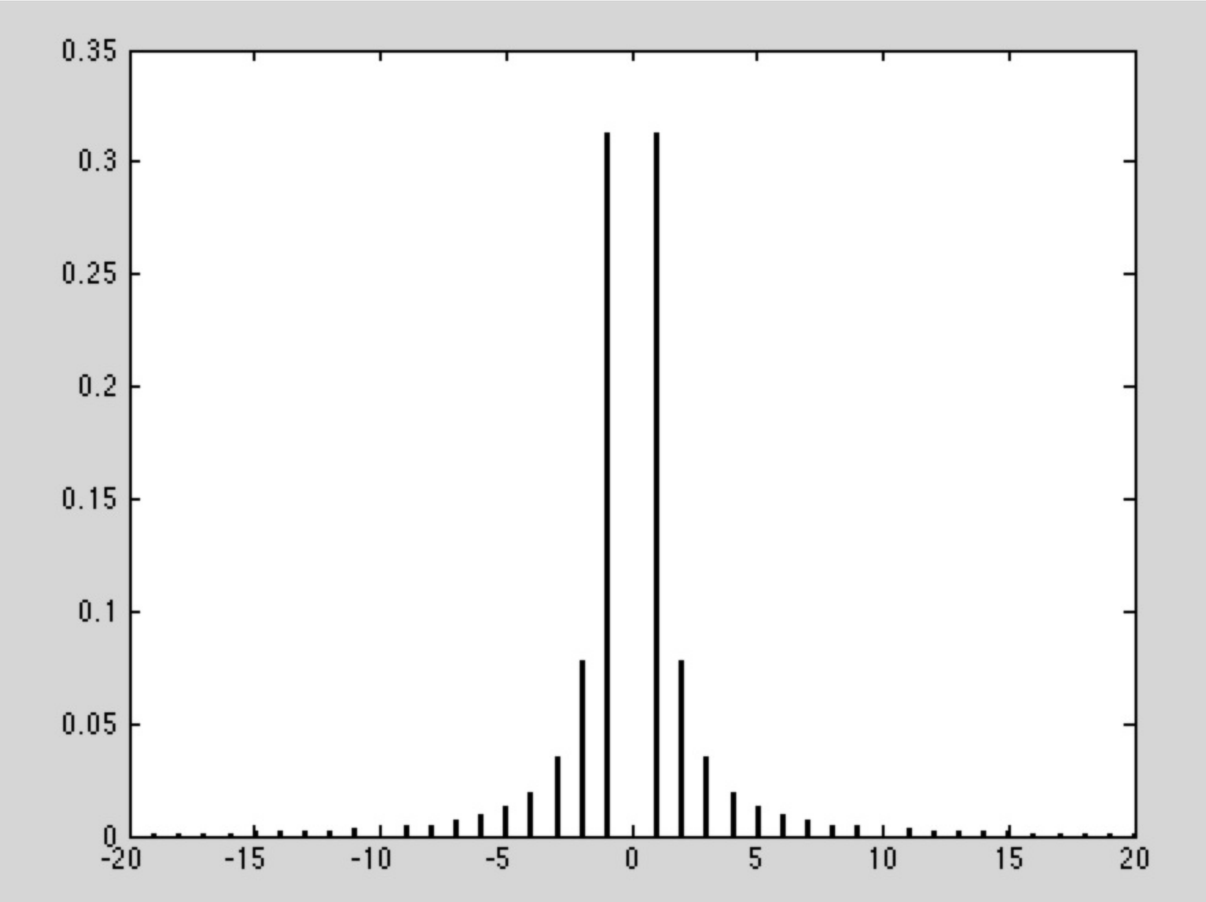
$$P(X=i)=\begin{cases} 0 & \text{if } i=0 \\ \dfrac{1}{Z|i|^3} & \text{if } i\neq 0 \end{cases} \quad ; \quad Z=2\sum_{i=1}^{\infty}\frac{1}{|i|^3}<\infty$$



$$E(X)=\sum_{i=1}^{\infty} iP(X=i)+\sum_{i=-1}^{-\infty} iP(X=i)=\frac{1}{Z}\left(\sum_{i=1}^{\infty}\frac{1}{i^2}-\sum_{i=1}^{\infty}\frac{1}{i^2}\right)=\frac{c-c}{Z}=0$$

A symmetric distribution on pos and neg integers, the bad case

$$P(X=i)=\begin{cases} 0 & \text{if } i=0 \\ \dfrac{1}{Zi^2} & \text{if } i\neq 0 \end{cases} \quad ; \quad Z=2\sum_{i=1}^{\infty}\frac{1}{i^2}<\infty$$

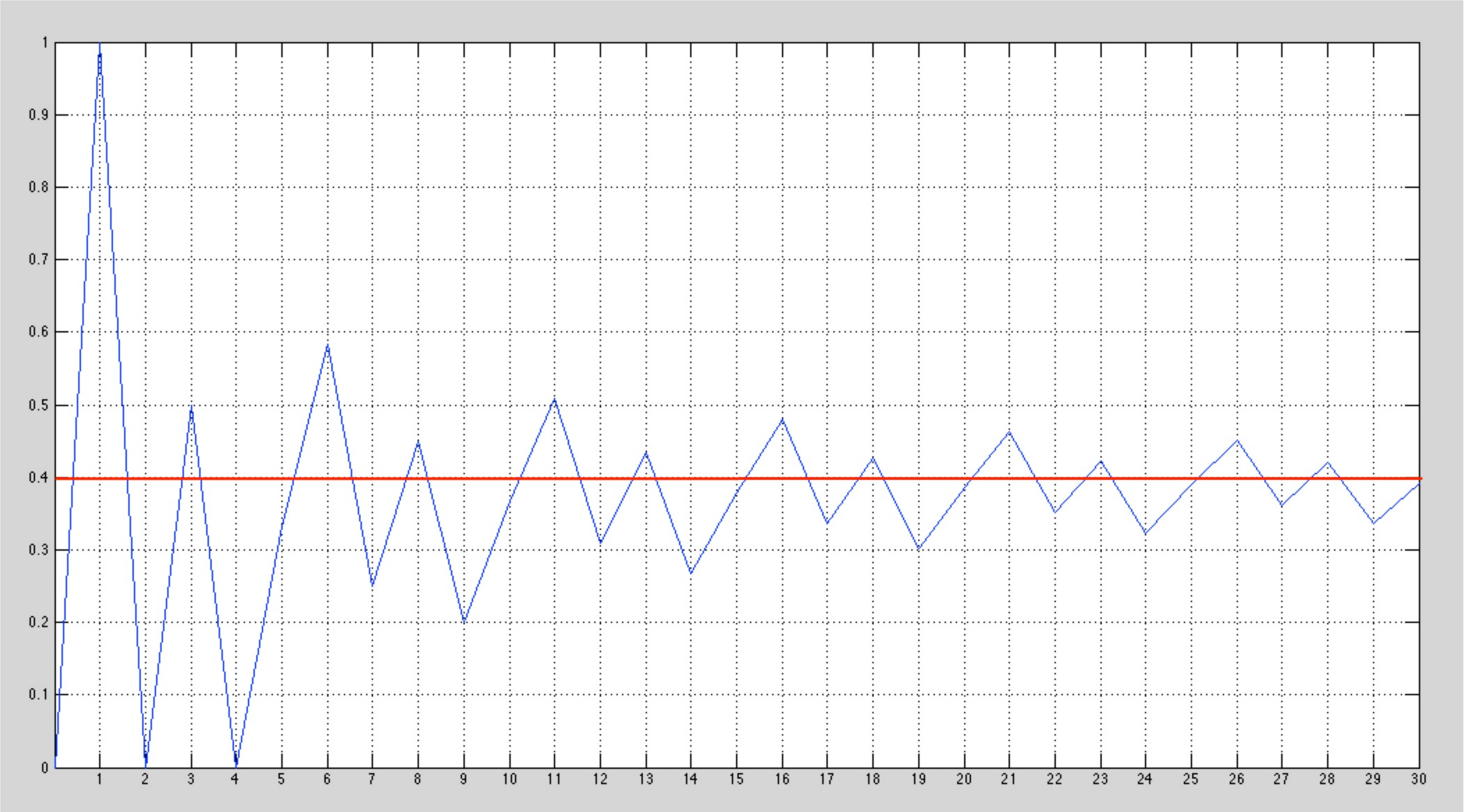


$$E(X)=\sum_{i=1}^{\infty} iP(X=i)+\sum_{i=-1}^{-\infty} iP(X=i)=\frac{1}{Z}\left(\sum_{i=1}^{\infty}\frac{1}{i}-\sum_{i=1}^{\infty}\frac{1}{i}\right)=\frac{\infty-\infty}{Z}=\mathit{undefined}$$

Undefined limit means you can get the limit of your choice by changing the order of summation.

You have at your disposal two infinitely large sums with shrinkingly small pieces:
1/1,1/2,1/3,1/4,....   -1/1,-1/2,-1/3,-1/4,...
Suppose you want the limit to be 0.4, by alternating between positives and negatives you can get arbitrarily close to 0.4 (or to any other number)

1/1-1/1+1/2-1/2+1/3+1/4-1/3+1/5-1/4+1/6+1/7-1/5+1/8-1/6+1/9+1/10-1/7
+1/11-1/8+1/12+1/13-1/9+1/14-1/10+1/15+1/16-1/11+1/17-1/12+1/18 = 0.3919



webwork.cse.ucsd.edu/misc/expected.html

Apps | Recent changes - Se | My Page | CSE103 | AWS | Googles | TWiki | UCSD

Let X be a random variable whose probability distribution is defined as:

$$P(X=i)=\begin{cases}\frac{1}{|i|^{\alpha}} & \text{if } i\neq 0\\ 0 & \text{if } i=0.\end{cases}$$

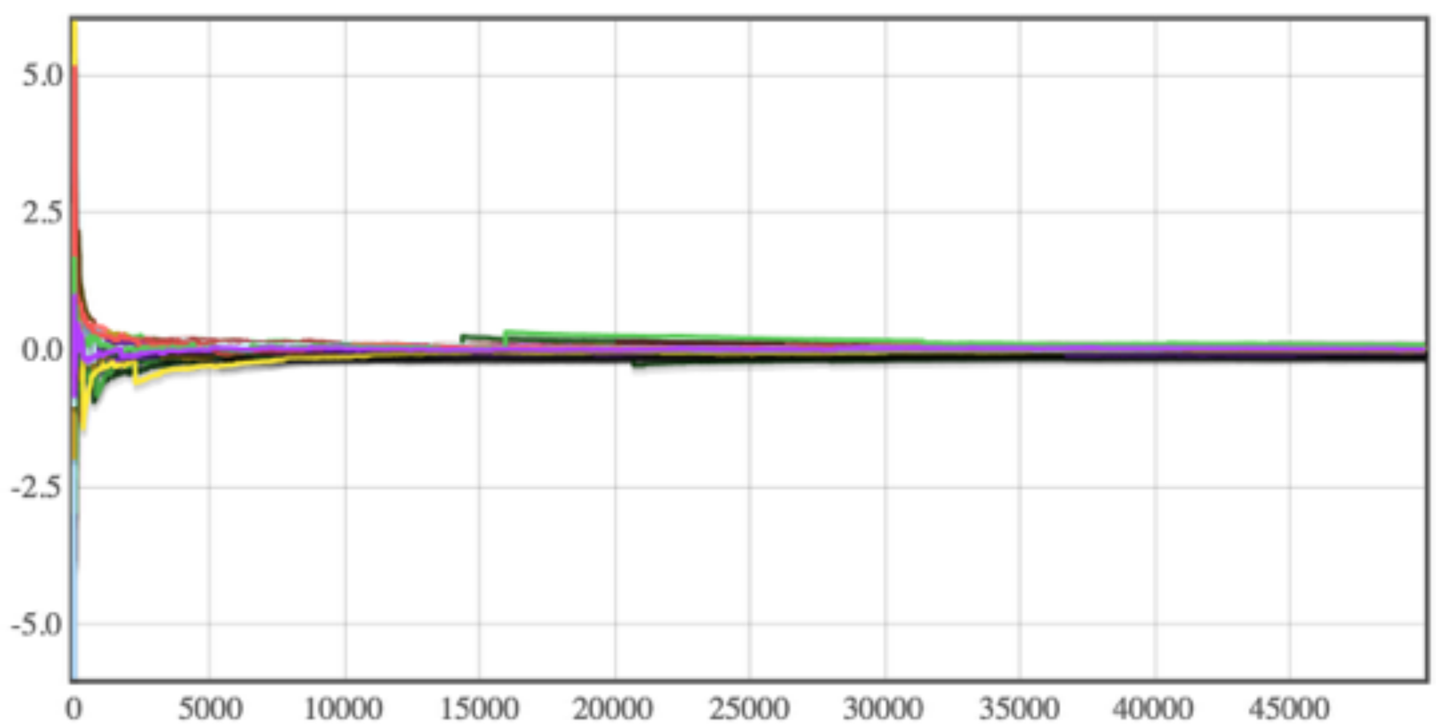
Simulation parameters:

α: 2.5

Number of trajectories: 50

Number of data points: 50000

Run



Let X be a random variable whose probability distribution is defined as:

$$P(X = i) = \begin{cases} \frac{1}{|i|^{\alpha}} & \text{if } i \neq 0 \\ 0 & \text{if } i = 0. \end{cases}$$

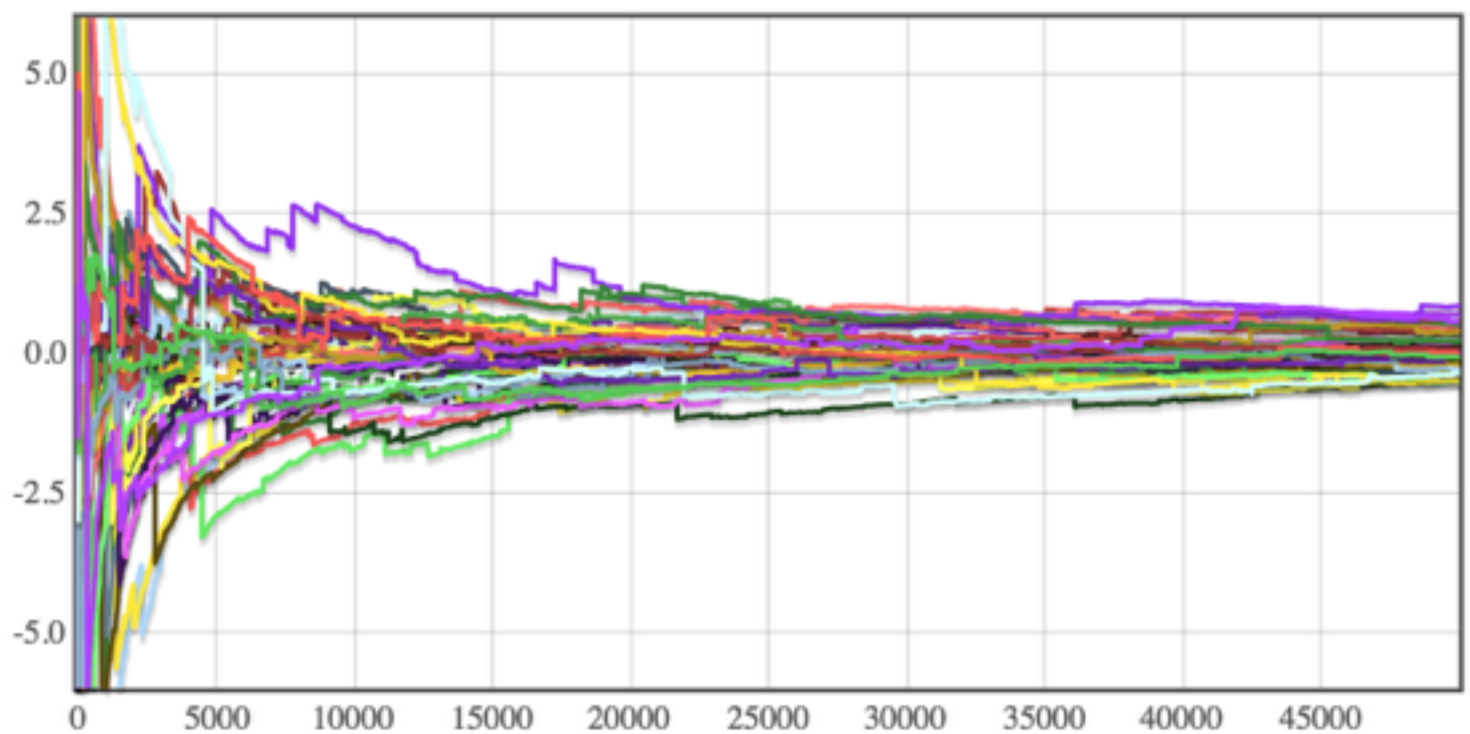
Simulation parameters:

α: 2.0

Number of trajectories: 50

Number of data points: 50000

Run



webwork.cse.ucsd.edu/misc/expected.html

Apps | Recent changes - Se | My Page | CSE103 | AWS | Googles | TWiki | UCSD

Let X be a random variable whose probability distribution is defined as:

$$P(X = i) = \begin{cases} \frac{1}{|i|^{\alpha}} & \text{if } i \neq 0 \\ 0 & \text{if } i = 0. \end{cases}$$

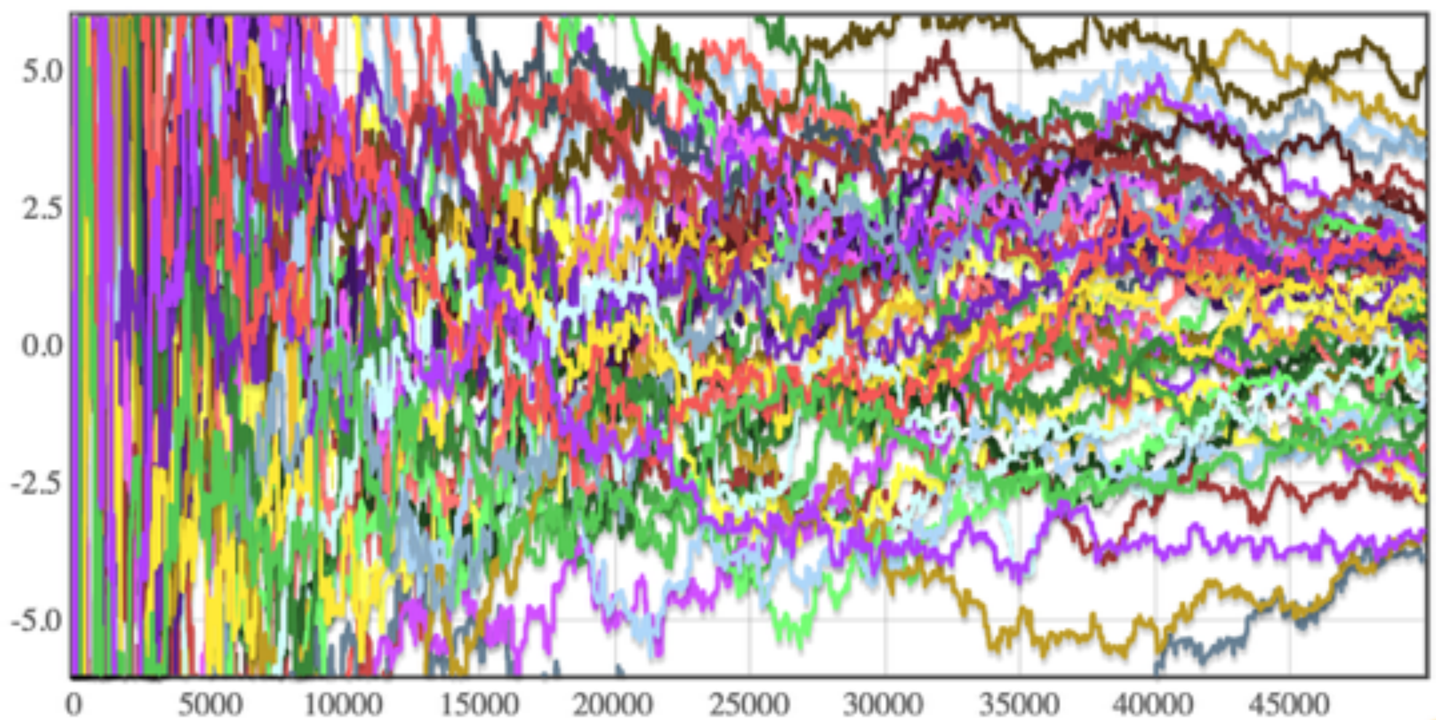
Simulation parameters:

α: 1.5

Number of trajectories: 50

Number of data points: 50000

Run



- Starting Next Time: Randomized algorithms
- Examples of harder problems using linearity of expectations will be done in discussion sessions.