Foundations of Data Science, Fall 2015, UC Berkeley

Example from Lecture 9/2/15

A. Adhikari

The Birthday Problem There are 120 students in a class. What is the chance that at least two of the students have the same birthday?

Assumptions

1. No leap years: Every year has 365 days.
2. No clumping of births at any time of year: Each student's birthday is equally likely to be any of the 365 days.
3. No twins, triplets, or other dependencies: No student's birthday is affected by any of the others.

In [8]:

```
# Answer to The Birthday Problem (with 4 people, not 120)

"""First find the chance that
all four people have different birthdays,
then subtract from 1.
Note that there are no restrictions
on the birthday of the first person you consider.
That is why there are only three factors in the product below."""

1 - (364/365)*(363/365)*(362/365)
```

Out[8]:

0.016355912466550215

In [10]:

```
# Now using arrays: The Birthday Problem (4 people)

# Create the array [364, 363, 362]

bdays4 = np.arange(364, 361, -1)
bdays4
```

Out[10]:

array([364, 363, 362])

```
In [11]:
```

```
# Divide each term by 365, to get [364/365, 363/365, 362/365]

bdays4_fractions = bdays4/365
bdays4_fractions
```

```
Out[11]:
```

```
array([ 0.99726027,  0.99452055,  0.99178082])
```

```
In [12]:
```

```
# Multiply them all together, and subtract from 1. Done!

bday_4_prob = 1 - np.prod(bdays4_fractions)
bday_4_prob
```

```
Out[12]:
```

```
0.016355912466550215
```

```
In [13]:
```

```
# Apply the same method to all class sizes between 2 and 365

bday_all = np.arange(364, 0, -1)
```
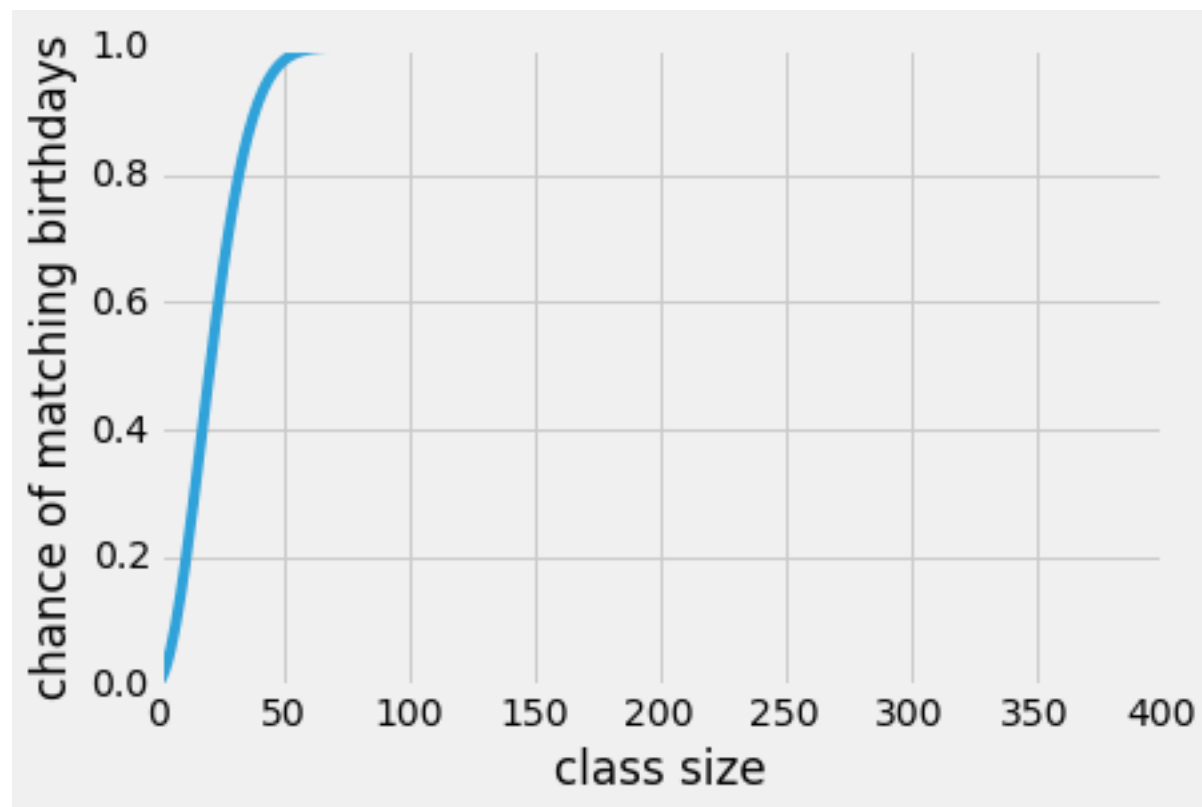
```
In [14]:
```

```
bday_probs = 1 - np.cumprod(bday_all/365)
```

```
In [15]:
```

```
plots.plot(bday_probs)
plots.ylabel("chance of matching birthdays")
plots.xlabel("class size")
```

```
Out[15]:
```

```
<matplotlib.text.Text at 0x1028b1438>
```



```
In [16]:
```

```
bday_probs
```

```
Out[16]:
```

```
array([ 0.00273973,  0.00820417,  0.01635591,  0.02713557,  0.040462
48,
        0.0562357 ,  0.07433529,  0.09462383,  0.11694818,  0.141141
38,
        0.16702479,  0.19441028,  0.22310251,  0.25290132,  0.283604
01,
        0.31500767,  0.34691142,  0.37911853,  0.41143838,  0.443688
34,
        0.47569531,  0.50729723,  0.53834426,  0.5686997 ,  0.598240
82,
        0.62685928,  0.65446147,  0.68096854,  0.70631624,  0.730454
63,
        0.75334753,  0.77497185,  0.79531686,  0.81438324,  0.832182
11,
        0.84873401,  0.86406782,  0.87821966,  0.89123181,  0.903151
61,
        0.91403047,  0.92392286,  0.93288537,  0.9409759 ,  0.948252
84,
        0.9547744 ,  0.96059797,  0.96577961,  0.97037358,  0.974431
99,
```

0.97800451, 0.98113811, 0.98387696, 0.98626229, 0.988332
35,
        0.99012246, 0.99166498, 0.99298945, 0.99412266, 0.995088
8 ,
        0.99590957, 0.99660439, 0.99719048, 0.99768311, 0.998095
7 ,
        0.99844004, 0.99872639, 0.99896367, 0.99915958, 0.999320
75,
        0.99945288, 0.99956081, 0.99964864, 0.99971988, 0.999777
44,
        0.99982378, 0.99986095, 0.99989067, 0.99991433, 0.999933
11,
        0.99994795, 0.99995965, 0.99996882, 0.999976 , 0.999981
59,
        0.99998593, 0.99998928, 0.99999186, 0.99999385, 0.999995
37,
        0.99999652, 0.9999974 , 0.99999806, 0.99999856, 0.999998
93,
        0.99999922, 0.99999942, 0.99999958, 0.99999969, 0.999999
78,
        0.99999984, 0.99999988, 0.99999992, 0.99999994, 0.999999
96,
        0.99999997, 0.99999998, 0.99999998, 0.99999999, 0.999999
99,
        0.99999999, 1.   , 1.   , 1.   , 1.
,
        1.   , 1.   , 1.   , 1.   , 1.
,
        1.   , 1.   , 1.   , 1.   , 1.
,
        1.   , 1.   , 1.   , 1.   , 1.
,
        1.   , 1.   , 1.   , 1.   , 1.
,
        1.   , 1.   , 1.   , 1.   , 1.
,
        1.   , 1.   , 1.   , 1.   , 1.
,
        1.   , 1.   , 1.   , 1.   , 1.
,
        1.   , 1.   , 1.   , 1.   , 1.
,
        1.   , 1.   , 1.   , 1.   , 1.
,
        1.   , 1.   , 1.   , 1.   , 1.
,
        1.   , 1.   , 1.   , 1.   , 1.
,
        1.   , 1.   , 1.   , 1.   , 1.
,
        1.   , 1.   , 1.   , 1.   , 1.

,

1.   , 1.   , 1.   , 1.   , 1.
,

1.   , 1.   , 1.   , 1.   , 1.
,

1.   , 1.   , 1.   , 1.   , 1.
,

1.   , 1.   , 1.   , 1.   , 1.
,

1.   , 1.   , 1.   , 1.   , 1.
,

1.   , 1.   , 1.   , 1.   , 1.
,

1.   , 1.   , 1.   , 1.   , 1.
,

1.   , 1.   , 1.   , 1.   , 1.
,

1.   , 1.   , 1.   , 1.   , 1.
,

1.   , 1.   , 1.   , 1.   , 1.
,

1.   , 1.   , 1.   , 1.   , 1.
,

1.   , 1.   , 1.   , 1.   , 1.
,

1.   , 1.   , 1.   , 1.   , 1.
,

1.   , 1.   , 1.   , 1.   , 1.
,

1.   , 1.   , 1.   , 1.   , 1.
,

1.   , 1.   , 1.   , 1.   , 1.
,

1.   , 1.   , 1.   , 1.   , 1.
,

1.   , 1.   , 1.   , 1.   , 1.
,

1.   , 1.   , 1.   , 1.   , 1.
,

1.   , 1.   , 1.   , 1.   , 1.
,

1.   , 1.   , 1.   , 1.   , 1.
,

1.   , 1.   , 1.   , 1.   , 1.
,

1.   , 1.   , 1.   , 1.   , 1.
,
          1.   , 1.   , 1.   , 1.   , 1.
,
          1.   , 1.   , 1.   , 1.   , 1.
,
          1.   , 1.   , 1.   , 1.   , 1.
,
          1.   , 1.   , 1.   , 1.   , 1.
,
          1.   , 1.   , 1.   , 1.   , 1.
,
          1.   , 1.   , 1.   , 1.   , 1.
,
          1.   , 1.   , 1.   , 1.   , 1.
,
          1.   , 1.   , 1.   , 1.   , 1.
,
          1.   , 1.   , 1.   , 1.   ])