

Distributions in Python

random variables are easy with `scipy.stats`

Clark Fitzgerald - @clarkfitzg

UC Davis - Statistics

24 November 2014

Topics

- Why I'm a bad programmer
- Introduction to scipy.stats (from Python!)
- Live Demo

Why I'm a bad programmer

- Error-prone
- Lazy

The Binomial Distribution

$$X \sim \text{Binom}(10, 0.5)$$

I'd like to know the probability mass function at 4.

Is this code correct? incorrect?

```
def f(n):  
    'computes n!'  
    return functools.reduce(operator.mul, range(n))  
  
def binom(n, k):  
    'computes n choose k'  
    return f(n) / (f(k) * f(n - k))  
  
def binom_pmf(n, p, k):  
    'The probability mass function of Binom(n, p)'  
    return binom(n, k) * p ** (n - k) * (1 - p) ** k
```

Is this code correct? incorrect?

Incorrect!

```
binom_pmf(10, 0.5, 4)
```

Thoughts

- algebra :(
- inefficient :(
- mistakes :(
- inconsistent :(
- maintenance :(

$$X \sim \text{Binom}(10, 0.5)$$

I'd like to know the probability mass function at 4.

Is this code correct? incorrect?

```
from scipy.stats import binom
```

```
X = binom(10, 0.5)
```

```
X.pmf(4)
```

Correct!

I want to program the way I think, and just have it work perfectly.

With RV's we ask the same questions over and over.

- What's the mean?
- Given some data, what's the MLE for a parameter?
- Give me n samples of some random variable X .
- What's the cdf of X with parameter β evaluated at k ?

That's why we had books of tables!

But now we have computers.

Discrete distributions

13 discrete distributions

bernoulli, binom, boltzmann, dlaplace, geom, hypergeom, logser, nbinom, planck, poisson, randint, skellam, zipf

Scipy Continuous Distributions

84 Continuous Distributions

alpha, anglit, arcsine, beta, betaprime, bradford, burr, cauchy, chi, chi2, cosine, dgamma, dweibull, erlang, expon, exponweib, exponpow, f, fatiguelife, fisk, foldcauchy, foldnorm, frechet, variable, frechet, genlogistic, genpareto, genexpon, genextreme, gausshyper, gamma, gengamma, genhalflogistic, gilbrat, gompertz, gumbel, halfcauchy, halflogistic, halfnorm, hypsecant, invgamma, invgauss, invweibull, johnsonsb, johnsonsu, ksone, kstwobign, laplace, logistic, loggamma, loglaplace, lognorm, lomax, maxwell, mielke, nakagami, ncx2, ncf, nct, norm, pareto, pearson3, powerlaw, powerlognorm, powernorm, rdist, reciprocal, rayleigh, rice, recipinvgauss, semicircular, t, triang, truncexpon, truncnorm, tukeylambda, uniform, vonmises, wald, weibull, variable, weibull, wrapcauchy