Working With Statistical Distributions

```
from scipy import stats
from scipy.stats import norm, binom
```

PDF / PMF

Probability our random variable takes on a given value.

- *pdf*: probability density function, for continuous distributions
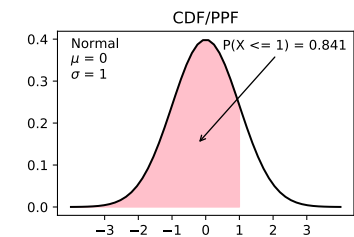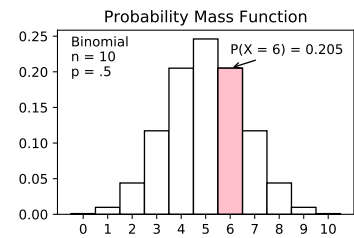- *pmf*: probability mass function, for discrete distributions

CDF & PPF

Probability our random variable takes on a value less than or equal to a given point.

- *cdf*: cumulative density function, given a value, what's the probability?
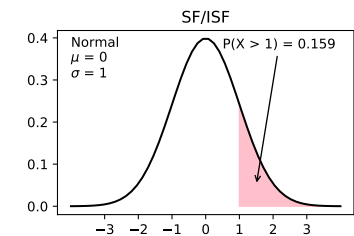- *ppf*: percent point function, given a probability, what's the value?

SF & ISF

Probability our random variable takes on a value greater than a given point.

- *sf*: survival function, given a value, what's the probability?
- *isf*: inverse survival function, given a probability, what's the value?

Scipy lets us specify parameters for the various distributions and then use the functions outlined above[1].

- A normal distribution with $\mu = 70, \sigma = 8$

  ```
  my_normal_distribution = norm(70, 8)
  ```

  - What's the likelihook the value is over 80?

    ```
    my_normal_distribution.sf(80)
    ```

  - What value is the cutoff point for the bottom 30%?

    ```
    my_normal_distribution.ppf(.3)
    ```

- A binomial distribution with $n = 10, p = .3$

  ```
  my_binomial_distribution = binom(10, .3)
  ```

  - What's the likelihood of 5 or less successes?

    ```
    my_binomial_distribution.cdf(5)
    ```

  - What's the probability of *exactly* 4 successes?

    ```
    my_binomial_distribution.pmf(4)
    ```


Percentile

[1] These parameters can also be specified as keyword arguments when the function is invoked, but for the sake of simplicity we will show this way.