Random Variables, Mean, and Variance

Random Variables

Definition: If X is a sample space, then a *random variable* is a real valued function $f : X \to \mathbf{R}$.

- Suppose that X is the sample space for a coin flip, so consists of heads and tails. Let $b(H) = 1$ and $b(T) = 0$. Then b is called a "Bernoulli Random Variable."
- For example, suppose that X is the set of N independent coin flips of a fair coin so that X consists of sequences of N heads or tails. If $x \in X$, let $f(x)$ be the number of heads. Then f is a random variable. Notice that f is the sum of N Bernoulli random variables.
- If X is a set of rolls of a pair of independent six-sided dice, and $f(x)$ is the sum of the values of the two dice, then f is another example of a random variable.
- If $U \subset \mathbf{R}$, and f is a random variable, then $f^{-1}(U) \subset X$ and, by definition,

$$P(f(x) \in U) = P(f^{-1}(U))$$

Example: For the coin-flipping example, suppose that $N = 4$ and f counts the number of heads. What is $P(f = 2)$?

Continuous random variable example

Suppose that $X = \mathbf{R}^2$ and

$$P(x \in U) = \left(\frac{1}{\sqrt{2\pi}}\right)^2 \int_U e^{-\|x\|^2/(2)} dx dy$$

Let $f(x) = \|x\|$. What is $P(f < r)$? In other words, how likely is a randomly drawn point to lie within distance r of the origin?

Independent random variables

Two random variables f and g are independent if:

- in the discrete case, $P(f = a \,\text{and}\, g = b) = P(f = a)P(g = b)$ for all $a, b \in \mathbf{R}$.
- in the continuous case, if $f^{-1}(U)$ and $g^{-1}(V)$ are independent for all intervals U and V in $\mathbf{R}$.

Expectation (Mean)

Definition: If f is a random variable on a sample space X, then

$$E[f] = \sum_{x \in X} f(x)P(x)$$

if X is discrete, or

$$E[f] = \int_X f(x)p(x)dx$$

where $p(x)$ is the density function, if X is continuous.

Properties:

- Linearity: $E[af + bg] = aE[f] + bE[g]$ if a and b are constants.
- If f and g are independent, then $E[fg] = E[f]E[g]$.

Example: If f is a binomial random variable with parameters N and p, then $E[f] = Np$.

Variance

Definition: If f is a random variable on a sample space X, then $\sigma^2(f)$, the *variance* of f is

$$\sigma^2(f) = E[(f - E[f])^2] = E[f^2] - E[f]^2$$

Variance of Binomial Random Variable

The variance of a binomial random variable with parameters N and p is $Np(1-p)$.

Variance of Normally Distributed Random Variable

The mean $E[x]$ of a normally distributed random variable with density

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/(2\sigma^2)}$$

is $E[x] = 0$.

The variance $E[x^2] - E[x]^2 = E[x^2] = \sigma^2$.