Discrete probability unit

- Random variables, introduction (last week Friday)
- Random variables, representing distributions
- Review (Wednesday)
- Test (Friday)
- Expected value and variance (**Today**)
- Common discrete distributions (Friday)
- Continuous random variables (next week)

Today:

- Definition and intuition of expected value
- Properties of expected value
- Variance

Given a sample space Ω, a **random variable** is a function whose domain is Ω. Usually random variables are numerically-valued, that is,

$$f : \Omega \to \mathbb{R}$$

Notationally, we use X, Y, etc for random variables and x, y, etc for values that a random variable can take on, and write

$$P(X = x) = P(\{a \in \Omega \mid X(a) = x\})$$

A random variable defines are partition of the sample space based on the results of a function computed from the basic outcomes. A random variable serves as an abstraction for the phenomenon being modeled.

The probability distribution for a random variable is a theoretical model for the empirical distribution of the data associated with a real population. If the model is an accurate representation of nature, the theoretical and empirical distributions are equivalent. Consequently, we attempt to acquire numerical descriptive measures, parameters, *for the probability distribution [to compare with similar measures on the data].*

Wackerly et al, Mathematical Statistics with Applications, 2008

For a random variable $X : \Omega \rightarrow \chi$ with probability mass function p, the **expected value** of X is

$$E[X] = \sum_{x \in \chi} x \; p(x)$$

1. Consider an urn with 14 balls: 8 white, 4 black, and 2 orange. You draw two balls randomly without replacement. For each black ball, you win \$2, and for each white ball we lose \$1 (orange balls don't count either way). Let random variable X denote your winnings. What is the expected value of X?
2. Consider rolling 3 fair, standard dice. How many *basic outcomes* are there? Let random variable X denote the values that can be rolled. What is the expected value of X?

Theorem 1

For any function f with domain χ,

$$E[f(X)] = \sum_{x \in \chi} f(x)\; p(x)$$

Theorem 2

For any $c \in \mathbb{R}$,

$$E[c] = \sum_{x \in \chi} c\; p(x) = c$$

Theorem 3

For any $a \in \mathbb{R}$,

$$E[aX] = \sum_{x \in \chi} ax\; p(x) = aE[X]$$

Theorem 4

For any $b \in \mathbb{R}$,

$$E[X + b] = \sum_{x \in \chi} (x + b)\; p(x) = E[X] + b$$

For random variable X with mean μ, the **variance** of X is

$$Var[X] = E\left[(X - \mu)^2\right]$$

Theorem 5

$Var(X) = E\left[X^2\right] - (E[X])^2$

Theorem 6

$Var(aX + b) = a^2 Var(X)$

Theorem 7

For random variable X with mean μ and variance σ^2,

$$E\left[(X - c)^2\right] = \sigma^2$$

For next time:

Do the exercises on the bottom of the handout.