Recap

- (Ch 4) Random variables
  - Probability distribution
  - Joint probability distribution
  - Conditional probability distribution

Today

- (Ch 4) Random variables
  - Expected value, variance and covariance
  - Towards the weak law of large numbers

Expected value

- The **expected value** of a random variable X is

$$E[X] = \sum_{x} xP(x)$$

- The expected value is a weighted average of the values taken by X

Expected value: gambling example

- Let's bet on a coin toss
  - It comes up heads with probability p and tails with probability $1 - p$
  - If it comes up heads I pay you $10; otherwise, you pay me $10
- For what values of p is this a good game for you?

Let X be your net winnings

$E[X] = 10p + (-10)(1-p) = 20p - 10$

For $E[X] > 0$, you must have $p > 0.5$

Expected value as mean

- Suppose we have a data set $\{x_i\}$ of N data points. Let's build an empirical probability distribution from the data set by assigning each data point with probability $^1/_N$.

$$E[X] = \sum_i x_i P(x_i) = \frac{1}{N} \sum_i x_i = \text{mean}(\{x_i\})$$

- The expected value is also called the **mean**

Linearity properties of expected value

- For random variables X and Y and constant k

$$E[X+Y] = E[X] + E[Y]$$

$$E[x+k] = E[x] + \underbrace{E[k]}_{k}$$

$$E[kX] = kE[X]$$

- These properties follow from interpreting expected values as means of data sets

Expected value of a function of X

- If f is a function of a random variable X, then $Y = f(X)$ is a random variable too

- The **expected value** of $Y = f(X)$ is

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

Expected value: online gambling example

Let's make the same bet as before, but now you pay a 10% fee on all transactions. For what values of p is this a good game for you?

X is your net winnings before fees. We already know that $E[X] = 20p - 10$

Let Y be your net winnings after fees, so $Y = X - 0.1|X|$

$$E[Y] = (10 - 0.1|10|)p + (-10 - 0.1|10|)(1-p)$$

$$= 9p + (-11)(1-p) = 20p - 11$$

For $E[Y] > 0$, you must have $p > 0.55$

Variance and standard deviation

- The **variance** of a random variable X is

$$\mathrm{var}[X] = E[(X - E[X])^2]$$

- The **standard deviation** of a random variable X is

$$\mathrm{std}[X] = \sqrt{\mathrm{var}[X]}$$

Properties of variance

- For random variables X and Y and constant k

$$\mathrm{var}[k] = 0$$

$$\mathrm{var}[X] \geq 0$$

$$\mathrm{var}[kX] = k^2 \mathrm{var}[X]$$

- If X and Y are independent

$$\mathrm{var}[X + Y] = \mathrm{var}[X] + \mathrm{var}[Y]$$

A neater expression for variance

$\mathrm{var}[X] = E[(X-\mu)^2]$ where $\mu = E[X]$

$= E[X^2 - 2X\mu + \mu^2]$

$= E[X^2] - 2\mu E[X] + \mu^2$ by linearity

$= E[X^2] - 2E[X]E[X] + (E[X])^2$

$= E[X^2] - (E[X])^2$

Variance: online gambling example

Let's make the same bet as before. What is the variance of your net winnings before fees?

X is your net winnings before fees. We already know that $E[X] = 20p - 10$

$$\text{var}[X] = E[X^2] - (E[X])^2$$

$$= \left(10^2 p + (-10)^2(1-p)\right) - (20p-10)^2$$

$$= 100 - (400p^2 - 400p + 100)$$

$$= 400p(1-p)$$

Covariance

- The **covariance** of random variables X and Y is

$$\text{cov}(X, Y) = E[(X - E[X])(Y - E[Y])]$$

- Note that

$$\text{cov}(X, X) = E[(X - E[X])^2] = \text{var}[X]$$

Properties of covariance

- A neater expression for covariance (similar derivation as for variance)

$$\text{cov}(X, Y) = E[XY] - E[X]E[Y]$$

- If X and Y are independent, the following are proven in the book

$$E[XY] = E[X]E[Y]$$

$$\text{cov}(X, Y) = 0$$

Covariance: online gambling example

Let's make the same bet as before. What is the covariance of your net winnings before fees and your net winnings after fees?

X and Y are your net winnings before and after fees. So, $E[X] = 20p - 10$ and $E[Y] = 20p - 11$

$$\text{cov}(X,Y) = E[XY] - E[X]E[Y]$$

$$= \left((10)(9)p + (-10)(-11)(1-p)\right) - (20p-10)(20p-11)$$

$$= (110 - 20p) - (400p^2 - 420p + 110)$$

$$= 400p(1-p)$$

Towards the weak law of large numbers

- The weak law says that if we repeat an experiment many times, the average of the observations will “converge” to the expected value

- For example, if you actually repeat the bet discussed in this lecture, your average winnings after fees will “converge” to $E[Y] = 20p - 11$

- The weak law justifies using simulations (instead of calculations) to estimate the expected values of random variables

Indicator functions

- An indicator function for an event E is a function of X such that

$$\mathbb{I}_{[E]}(x) = \begin{cases} 1 & \text{if } E \text{ occurs for this value of } X \\ 0 & \text{if } E \text{ does not occur for this value of } X \end{cases}$$

- The expected value of the indicator function is the probability of E

$$E\left[\mathbb{I}_{[E]}(x)\right] = 1 \cdot P(E) + 0\,(1 - P(E)) = P(E) \quad \text{— } (\star)$$

Markov's inequality

- For any random variable X and constant $a > 0$

$$P(|X| \geq a) \leq \frac{E[|X|]}{a}$$

- In words, a random variable is unlikely to have an absolute value much larger than the mean of its absolute value

- For example, if $a = 10E[|X|]$

$$P(|X| \geq 10E[|X|]) \leq 0.1$$

Proof of Markov's inequality

$$\mathbb{I}_{[|x| \geq a]}(x) = \begin{cases} 1 & \text{if } |x| \geq a \\ 0 & \text{if } |x| < a \end{cases}$$

$$\leq \frac{|x|}{a}$$

$$\underbrace{E\left[\mathbb{I}_{[|x| \geq a]}(x)\right]}_{P(|X| \geq a)} \leq \frac{E[|x|]}{a}$$

because of ✪

Chebyshev's inequality

- For any random variable X and constant $a > 0$

$$P(|X - E[X]| \geq a) \leq \frac{\mathrm{var}[X]}{a^2}$$

- To rephrase, let $a = k\sigma$ where $\sigma = \mathrm{std}[X]$

$$P(|X - E[X]| \geq k\sigma) \leq \frac{1}{k^2}$$

- In words, the probability that X is greater than k standard deviations from the mean is small