COMPSCI 240: Reasoning Under Uncertainty

Andrew Lan and Nic Herndon

University of Massachusetts at Amherst

Spring 2019

Lecture 10: Expectation and Variance

Recap: Common Discrete Random Variables

- **Uniform:** For $k = a, \ldots, b$:

$$P(X = k) = \frac{1}{b - a + 1}$$

- **Bernoulli:** For $k = 0$ or 1:

$$P(X = k) = \begin{cases} 1 - p & \text{if } k = 0 \\ p & \text{if } k = 1 \end{cases}$$

- **Binomial:** For $k = 0, \ldots, N$

$$P(X = k) = \binom{N}{k} p^k (1 - p)^{N-k}$$

- **Geometric:** For $k = 1, 2, 3, \ldots$, $P(X = k) = (1 - p)^{k-1} \cdot p$
- **Poisson:** $P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$ for $k = 0, 1, 2, \ldots$

Expected Value

- For a random variable X, the expected value is defined to be:

$$E[X] = \sum_{x \in \mathbb{R}} x \, P(X = x)$$

  i.e., the probability-weighted average of the possible values of X.
- $E[X]$ is also called the **expectation** or **mean** of X.
- Why do we care to know about the expected value?
- Given a certain PMF, what is the "average" outcome that I am expecting to have?
- For example, if I bet the same amount of money on roulette and play it for a long-term period, how much do I expect to make?
- For a long-term period, can you make money from casino?

Expected Value: Question

- Expectation:

$$E[X] = \sum_{k \in \mathbb{R}} k\, P(X = k)$$

- If X maps to $\{1, 2, 6\}$ and

$$P(X = 1) = 1/3 \quad , \quad P(X = 2) = 1/2 \quad , \quad P(X = 6) = 1/6$$

then is the expectation:

$A)$ 2 $\quad B)$ 2.33... $\quad C)$ 3 $\quad D)$ 3.5 $\quad E)$ 3.66...

- Answer is $E[X] = 1 \cdot \frac{1}{3} + 2 \cdot \frac{1}{2} + 6 \cdot \frac{1}{6} = 2.33\ldots$

Example: Expected Winnings in Games of Chance

- Suppose you play a simple game with your friend where you flip a coin. If the coin is heads, your friend pays you a dollar. If it's tails, you pay your friend a dollar.
- In any game of chance like this, you might be interested in how much money you might win or lose per round on average if you played many rounds.
- Suppose you play N rounds of the game and you win N_W times and lose N_L times. Your average payoff would be:

$$\frac{N_L(-1) + N_W(1)}{N}$$

Example: Expected Winnings in Games of Chance

- As the number of rounds increases, you will see that N_W/N converges to the probability of heads p, while N_L/N converges to the probability of tails $(1-p)$.
- Let X be a random variable mapping the outcomes $\{H, T\}$ to the payoff values $\{1, -1\}$. The limiting value of your average payoff converges to a number called the **expected value** of the random variable X:

$$E[X] = P(X = -1)(-1) + P(X = 1)(1)$$

- If the coin is fair, your expected winnings are:

$$\begin{aligned} E[X] &= P(X = -1)(-1) + P(X = 1)(1) \\ &= (0.5)(-1) + (0.5)(1) = 0 \end{aligned}$$

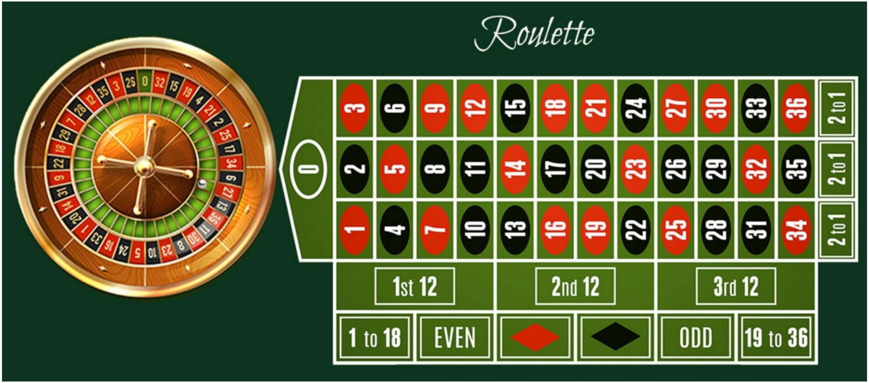
Example: Expected Winnings in Games of Chance

- Suppose your friend decides to trick you and swaps the fair coin for a biased coin that comes up tails with probability 0.7. How does this change your expected winnings?
- Your expected winnings are now computed as follows:

$$\begin{aligned} E[X] &= P(X = -1)(-1) + P(X = 1)(1) \\ &= (0.7)(-1) + (0.3)(1) \\ &= -0.4 \end{aligned}$$

- The interpretation is that over many many rounds of play with the biased coin, you would expect to lose forty cents per round on average.

Can You Make Money from Roulette?



One Challenging Problem on Expectation

We randomly pick 3 numbers from 10 integer numbers 1, 2, 3, 4, ...,10.
If the largest number among the 3 picked out is denoted as L, what is the expected value of L?
Answer:

$$P(X = L) = \begin{cases} 0 & \text{if } L < 3 \\ \frac{1 \cdot \binom{L-1}{2}}{\binom{10}{3}} & \text{if } L \geq 3 \end{cases}$$

$$\begin{aligned} E(X) &= \sum_{k=3}^{10} k \cdot P(X = k) = \sum_{k=3}^{10} k \frac{\binom{k-1}{2}}{\binom{10}{3}} \\ &= \sum_{k=3}^{10} k \frac{\frac{(k-1)!}{(k-1-2)! \cdot 2!}}{\binom{10}{3}} = \sum_{k=3}^{10} \frac{k(k-1)(k-2)}{2 \cdot \binom{10}{3}} \\ &= 8.25 \end{aligned}$$

Expectations of Common Random Variables

- **Uniform on** $\{a, a+1, \ldots, b\}$**:** $E[X] = \frac{a+b}{2}$
- **Bernoulli:** $E[X] = (1-p) \cdot 0 + p \cdot 1 = p$
- **Binomial:** $E[X] = \sum_{k=0}^{n} k \cdot \binom{n}{k} p^k (1-p)^{n-k} = np$
- **Geometric:** $E[X] = \sum_{k=1}^{\infty} k \cdot (1-p)^{k-1} p = \frac{1}{p}$
- **Poisson:** $E[X] = \sum_{k=0}^{\infty} k \cdot \frac{e^{-\lambda}}{k!} \lambda^k = \lambda$

Uniform Expectation

$$
\begin{aligned}
E[X] &= \sum_{k=a}^{b} kP(X = k) \\
&= \sum_{k=a}^{b} k \cdot \frac{1}{b - a + 1} \\
&= \frac{1}{b - a + 1} \sum_{k=a}^{b} k \\
&= \frac{1}{b - a + 1} \cdot \frac{(a + b)(b - a + 1)}{2} \\
&= \frac{a + b}{2}
\end{aligned}
$$

Binomial Expectation

$$
\begin{aligned}
E[X] &= \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} \cdot k \\
&= 0 + \sum_{k=1}^{n} \binom{n}{k} p^k (1-p)^{n-k} \cdot k \\
&= \sum_{k=1}^{n} \frac{n(n-1)\cdots(n-k+1)}{k!} p^k (1-p)^{n-k} \cdot k \\
&= np \sum_{k=1}^{n} \frac{(n-1)\cdots(n-k+1)}{(k-1)!} p^{(k-1)} (1-p)^{n-k} \quad \text{...let } l = k-1 \text{ and } m = n-1 \\
&= np \sum_{l=0}^{m} \frac{m\cdots(m-l+1)}{(l)!} p^l (1-p)^{m-l} \\
&= np \sum_{l=0}^{m} \frac{m!}{(l)!(m-l)!} p^l (1-p)^{m-l} \\
&= np \cdot 1 = np
\end{aligned}
$$

Geometric Expectation

$$
\begin{aligned}
E[X] &= \sum_{k=1}^{\infty} k \cdot (1-p)^{k-1} p \\
&= \sum_{k=1}^{\infty} (1-p)^{k-1} p + \sum_{k=1}^{\infty} (k-1) \cdot (1-p)^{k-1} p \\
&= 1 + (1-p) \sum_{k=2}^{\infty} (k-1) \cdot (1-p)^{k-2} p \\
&= 1 + (1-p)(1 \cdot P(X=1) + 2 \cdot P(X=2) + 3 \cdot P(X=3) + \ldots) \\
&= 1 + (1-p) E[X]
\end{aligned}
$$

and so $E[X] = 1/p$.

Poisson Expectation

$$E[X] = \sum_{k=0}^{\infty} k \cdot \frac{e^{-\lambda}}{k!}\lambda^k = 0 + \sum_{k=1}^{\infty} k \cdot \frac{e^{-\lambda}}{k!}\lambda^k$$

$$= \sum_{k=1}^{\infty} \frac{e^{-\lambda}}{(k-1)!}\lambda^k$$

$$= \sum_{k=1}^{\infty} \frac{e^{-\lambda}}{(k-1)!}\lambda^{(k-1)} \cdot \lambda$$

$$= \lambda \sum_{k=1}^{\infty} \frac{e^{-\lambda}}{(k-1)!}\lambda^{k-1}$$

Let $m = k - 1$

$$= \lambda \sum_{m=0}^{\infty} \frac{e^{-\lambda}}{(m)!}\lambda^m$$

$$= \lambda \cdot (P(X = 0) + P(X = 1) + P(X = 2) + \dots)$$

$$= \lambda$$

Properties of Expectation

- **Linearity of Expectation:** If a and b are any real values, then the expectation of $aX + b$ is:

$$E[aX + b] = a \cdot E[X] + b$$

- **Expectation of Expectation:** Applying the expectation operator more than once has no effect. $E[E[X]] = E[X]$ since $E[X]$ is already a constant.

Variance

- **Definition**: Variance measures how far we expect a random variable to be from its average.
- It measures the expectation of the squared deviation of a random variable from its mean.

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

- An equivalent definition is

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

(Proof?)

Variance

- **Definition**: Variance measures how far we expect a random variable to be from its average:

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

- An equivalent definition is

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

- **Definition**: we generally define the $\mathbf{n^{th}}$ **moment** of X as $E[X^n]$, the expected value of the random variable X^n.

Example 1

- Consider a random variable X_1 where

$$P(X_1 = 2) = 1/2 \qquad P(X_1 = 3) = 1/4 \qquad P(X_1 = 5) = 1/4$$

- The expected value is:

$$E[X_1] = \frac{1}{2} \cdot 2 + \frac{1}{4} \cdot 3 + \frac{1}{4} \cdot 5 = 3$$

- The variance is:

$$\text{var}[X_1] = \frac{1}{2} \cdot (2-3)^2 + \frac{1}{4} \cdot (3-3)^2 + \frac{1}{4} \cdot (5-3)^2 = 1.5$$

Example 2

- Consider a random variable X_2 where

$$P(X_2 = -1) = 1/2 \qquad P(X_2 = 7) = 1/2$$

- The expected value is:

$$E[X_2] = \frac{1}{2} \cdot (-1) + \frac{1}{2} \cdot 7 = 3$$

- The variance is:

$$\text{var}[X_2] = \frac{1}{2} \cdot (-1 - 3)^2 + \frac{1}{2} \cdot (7 - 3)^2 = 16$$

Example 1 and 2

- Both examples shared the same expected value:

$$E[X_1] = \frac{1}{2} \cdot 2 + \frac{1}{4} \cdot 3 + \frac{1}{4} \cdot 5 = 3$$

$$E[X_2] = \frac{1}{2} \cdot (-1) + \frac{1}{2} \cdot 7 = 3$$

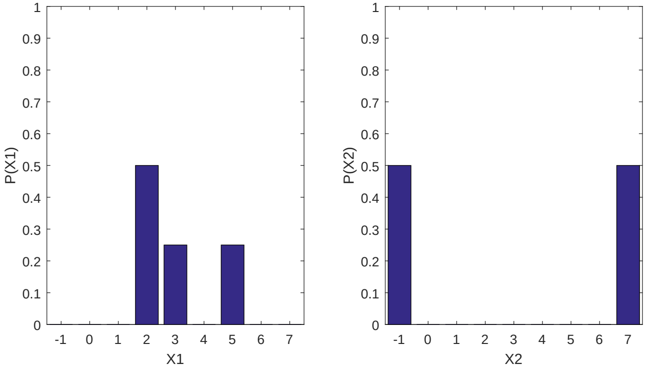
- But the variances were different:

$$\text{var}[X_1] = \frac{1}{2} \cdot (2-3)^2 + \frac{1}{4} \cdot (3-3)^2 + \frac{1}{4} \cdot (5-3)^2 = 1.5$$

$$\text{var}[X_2] = \frac{1}{2} \cdot (-1-3)^2 + \frac{1}{2} \cdot (7-3)^2 = 16$$

- What does this tell us?

Example 1 and 2



- We previously said that variance measures how far we expect a random variable to be from its average value.
- In other words, it measures **how spread the PMF looks like** with respect to the mean value.
- Why do we care about variance?