

Discrete Random Variables: Expectation, Mean and Variance

Berlin Chen

Department of Computer Science & Information Engineering
National Taiwan Normal University

Reference:

- D. P. Bertsekas, J. N. Tsitsiklis, *Introduction to Probability* , Sections 2.3-2.4

Motivation (1/2)

- **An Illustrative Example:** Suppose that you spin the wheel k times, and that k_i is the number of times that the outcome (money) is m_i (there are n distinct outcomes, m_1, m_2, \dots, m_n)
- What is the amount of money that you “*expect*” to get “*per spin*”?
 - The total amount received is

$$m_1 k_1 + m_2 k_2 + \dots + m_n k_n$$

- The amount received per spin is

$$M = \frac{m_1 k_1 + m_2 k_2 + \dots + m_n k_n}{k}$$

Motivation (2/2)

- If the number of spins k is very large, and if we are willing to interpret probabilities as relative frequencies, it is reasonable to anticipate that m_i comes up a fraction of times that is roughly equal to p_i

$$p_i \approx \frac{k_i}{k}$$

- Therefore, the amount received per spin can be also represented as

$$\begin{aligned} M &= \frac{m_1 k_1 + m_2 k_2 + \cdots + m_n k_n}{k} \\ &= m_1 p_1 + m_2 p_2 + \cdots + m_n p_n \end{aligned}$$

Expectation

- The **expected value** (also called the **expectation** or the **mean**) of a random variable X , with PMF p_X , is defined by

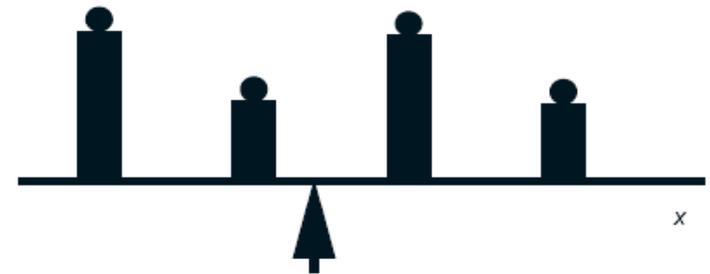
$$\mathbf{E}[X] = \sum_x xp_X(x)$$

- Can be interpreted as the **center of gravity** of the PMF (Or a weighted average, in proportion to probabilities, of the possible values of X)

- The expectation is well-defined if

$$\sum_x |x|p_X(x) < \infty$$

- That is, $\sum_x xp_X(x)$ converges to a finite value



Center of Gravity
c = Mean $E[X]$

$$\sum_x (x-c)p_X(x) = 0$$

$$\Rightarrow c = \sum_x x \cdot p_X(x)$$

Moments

- The **n -th moment** of a random variable X is the expected value of a random variable X^n (or the random variable Y , $Y = g(X) = X^n$)

$$\mathbf{E} \left[X^n \right] = \sum_x x^n p_X(x)$$

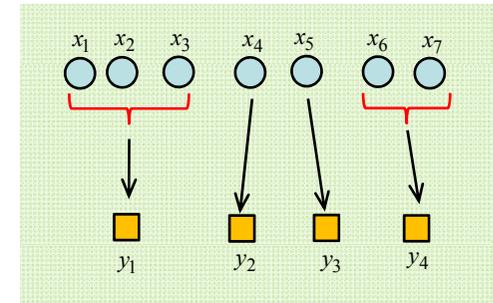
- The 1st moment of a random variable X is just its mean (or expectation)

X^n is termed as X raised to the power of n (or the n th power), or the n th power of X .

Expectations for Functions of Random Variables

- Let X be a random variable with PMF p_X , and let $g(X)$ be a function of X . Then, the expected value of the random variable $g(X)$ is given by

$$\mathbf{E}[g(X)] = \sum_x g(x) p_X(x)$$



- To verify the above rule

- Let $Y = g(X)$, and therefore $p_Y(y) = \sum_{\{x|g(x)=y\}} p_X(x)$

$$\mathbf{E}[g(X)] = \mathbf{E}[Y] = \sum_y y p_Y(y)$$

$$= \sum_y y \sum_{\{x|g(x)=y\}} p_X(x) = \sum_y \sum_{\{x|g(x)=y\}} g(x) p_X(x)$$

$$= \sum_x g(x) p_X(x)$$

?

Variance

- The **variance** of a random variable X is the expected value of a random variable $(X - \mathbf{E}(X))^2$

$$\begin{aligned}\text{var}(X) &= \mathbf{E} \left[(X - \mathbf{E}[X])^2 \right] \\ &= \sum_x (x - \mathbf{E}[X])^2 p_X(x)\end{aligned}$$

- The variance is always nonnegative (why?)
- The variance provides a measure of dispersion of X around its mean
- The standard derivation is another measure of dispersion, which is defined as (a square root of variance)

$$\sigma_X = \sqrt{\text{var}(X)}$$

- Easier to interpret, because it has the same units as X

An Example

- **Example 2.3:** For the random variable X with PMF

$$p_X(x) = \begin{cases} 1/9, & \text{if } x \text{ is an integer in the range } [-4, 4], \\ 0, & \text{otherwise} \end{cases} \quad \text{Discrete Uniform Random Variable}$$

$$\mathbf{E}[X] = \sum_x x p_X(x) = \frac{1}{9} \sum_{x=-4}^4 x = 0$$

$$\text{var}(X) = \mathbf{E}\left[(X - \mathbf{E}[X])^2\right] = \sum_x (x - \mathbf{E}[X])^2 p_X(x) = \frac{1}{9} \sum_{x=-4}^4 x^2 = \frac{60}{9}$$

$$\text{Or, let } Z = (X - \mathbf{E}[X])^2 = X^2$$

$$\Rightarrow p_Z(z) = \begin{cases} 2/9, & \text{if } z = 1, 4, 9, 16 \\ 1/9, & \text{if } z = 0 \\ 0, & \text{otherwise} \end{cases}$$

$$\text{var}(X) = \mathbf{E}[Z] = \sum_z z p_Z(z) = \frac{60}{9}$$

Properties of Mean and Variance (1/2)

- Let X be a random variable and let

$$Y = aX + b \quad \text{a linear function of } X$$

where a and b are given scalars

Then,

$$\mathbf{E}[Y] = a\mathbf{E}[X] + b$$

$$\text{var}(Y) = a^2 \text{var}(X)$$

- If $g(X)$ is a linear function of X , then

$$\mathbf{E}[g(X)] = g(\mathbf{E}[X]) \quad \text{How to verify it?}$$

Properties of Mean and Variance (2/2)

$$\mathbf{E}[Y] = \sum_x (ax + b)p_X(x) = \left[a \sum_x xp_X(x) \right] + \left[\underbrace{b \sum_x p_X(x)}_1 \right] = a\mathbf{E}[X] + b$$

$$\begin{aligned} \text{var}(Y) &= \sum_x (ax + b - \mathbf{E}[aX + b])^2 p_X(x) \\ &= \sum_x (ax + b - a\mathbf{E}[X] - b)^2 p_X(x) \\ &= a^2 \sum_x (x - \mathbf{E}[X])^2 p_X(x) \\ &= a^2 \underbrace{\sum_x (x - \mathbf{E}[X])^2 p_X(x)}_{\text{var}(X)} \\ &= a^2 \text{var}(X) \end{aligned}$$

Variance in Terms of Moments Expression

- We can also express variance of a random variable X as

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

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

$$= \sum_x (x^2 - 2x\mathbf{E}[X] + (\mathbf{E}[X])^2) p_X(x)$$

$$= \left[\sum_x x^2 p_X(x) \right] + 2\mathbf{E}[X] \left[\sum_x x p_X(x) \right] + (\mathbf{E}[X])^2 \left[\sum_x p_X(x) \right]$$

$$= \mathbf{E}[X^2] - 2(\mathbf{E}[X])^2 + (\mathbf{E}[X])^2$$

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

An Example

- **Example 2.4: Average Speed Versus Average Time.**

If the weather is good (with probability 0.6), Alice walks the 2 miles to class at a speed of $V=5$ miles per hour, and otherwise rides her motorcycle at a speed of $V=30$ miles per hour. What is the expected time $\mathbf{E}[T]$ to get to the class ?

$$p_V(v) = \begin{cases} 0.6, & \text{if } v = 5 \\ 0.4, & \text{if } v = 30 \end{cases}$$

$$\mathbf{E}[V] = 0.6 \times 5 + 0.4 \times 30 = 15$$

$$T = g(V) = \frac{2}{V}$$

$$\mathbf{E}[T] = 0.6 \times \frac{2}{5} + 0.4 \times \frac{2}{30} = \frac{4}{15}$$

$$\Rightarrow p_T(t) = \begin{cases} 0.6, & \text{if } t = \frac{2}{5} \\ 0.4, & \text{if } t = \frac{2}{30} \end{cases}$$

$$\text{However, } \mathbf{E}[T] = \mathbf{E}[g(V)] \neq g(\mathbf{E}[V]) = \frac{2}{15}$$

Mean and Variance of the **Bernoulli**

- **Example 2.5.** Consider the experiment of tossing a biased coin, which comes up a head with probability p and a tail with probability $1 - p$, and the **Bernoulli** random variable with PMF

$$p_X(x) = \begin{cases} p, & \text{if } x = 1 \\ 1 - p, & \text{if } x = 0 \end{cases}$$

$$\mathbf{E}[X] = \sum_x xp_X(x) = 1 \cdot p + 0 \cdot (1 - p) = p$$

$$\mathbf{E}[X^2] = \sum_x x^2 p_X(x) = 1^2 \cdot p + 0^2 \cdot (1 - p) = p$$

$$\text{var}(X) = \mathbf{E}[X^2] - (\mathbf{E}[X])^2 = p - p^2 = p(1 - p)$$

Mean and Variance of the **Discrete Uniform**

- Consider a **discrete uniform** random variable with a nonzero PMF in the range $[a, b]$

$$p_X(x) = \begin{cases} \frac{1}{b-a+1}, & \text{if } x = a, a+1, \dots, b \\ 0, & \text{otherwise} \end{cases}$$

$$\mathbf{E}[X] = \sum_x xp_X(x) = \frac{1}{b-a+1} \sum_{x=a}^b x = \frac{a+b}{2}$$

$$\mathbf{E}[X^2] = \frac{1}{b-a+1} \sum_{x=a}^b x^2 = \frac{1}{b-a+1} \cdot \left(\frac{b(b+1)(2b+1)}{6} - \frac{(a-1)(a)(2a-1)}{6} \right)$$

$$\begin{aligned} \text{var}(X) &= \mathbf{E}[X^2] - (\mathbf{E}[X])^2 = \frac{1}{b-a+1} \cdot \left(\frac{b(b+1)(2b+1)}{6} - \frac{(a-1)(a)(2a-1)}{6} \right) - \left(\frac{a+b}{2} \right)^2 \\ &= \frac{1}{b-a+1} \cdot \frac{(b-a)(b-a+1)(b-a+2)}{12} = \frac{(b-a)(b-a+2)}{12} \end{aligned}$$

Mean and Variance of the **Poisson**

- Consider a **Poisson** random variable with a PMF

$$p_X(x) = e^{-\lambda} \frac{\lambda^x}{x!}, \quad x = 0, 1, 2, \dots,$$

$$\mathbf{E}[X] = \sum_x x p_X(x) = \sum_{x=0}^{\infty} x e^{-\lambda} \frac{\lambda^x}{x!} = \lambda \sum_{x=1}^{\infty} e^{-\lambda} \frac{\lambda^{x-1}}{(x-1)!} = \lambda \sum_{x'=0}^{\infty} e^{-\lambda} \frac{\lambda^{x'}}{x'!} = \lambda$$

1

$$\begin{aligned} \mathbf{E}[X^2] &= \sum_x x^2 p_X(x) = \sum_{x=0}^{\infty} x^2 e^{-\lambda} \frac{\lambda^x}{x!} = \lambda \sum_{x=1}^{\infty} x \frac{\lambda^{x-1}}{(x-1)!} \\ &= \lambda \sum_{x'=0}^{\infty} (x'+1) e^{-\lambda} \frac{\lambda^{x'}}{x'!} = \lambda \left[\left(\sum_{x'=0}^{\infty} x' e^{-\lambda} \frac{\lambda^{x'}}{x'!} \right) + \left(\sum_{x'=0}^{\infty} e^{-\lambda} \frac{\lambda^{x'}}{x'!} \right) \right] = \lambda(\mathbf{E}[X] + 1) = \lambda^2 + \lambda \end{aligned}$$

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

1

Mean and Variance of the **Binomial**

- Consider a **binomial** random variable with a PMF

$$p_X(x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, \dots, n$$

$$\begin{aligned} \mathbf{E}[X] &= \sum_x x p_X(x) = \sum_{x=0}^n x \binom{n}{x} p^x (1-p)^{n-x} = \sum_{x=1}^n x \binom{n}{x} p^x (1-p)^{n-x} = \sum_{x=1}^n x \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \\ &= np \sum_{x=1}^n \frac{(n-1)!}{(x-1)!(n-x)!} p^{x-1} (1-p)^{n-x} = np \sum_{x'=0}^{n-1} \frac{(n-1)!}{x'!(n-1-x')!} p^{x'} (1-p)^{n-1-x'} = np \end{aligned}$$

$$\mathbf{E}[X^2] = \mathbf{E}[X^2 - X] + \mathbf{E}[X] \quad \text{(to be verified later on!)} \quad \mathbf{1}$$

$$\begin{aligned} \mathbf{E}[X^2 - X] &= \mathbf{E}[X(X-1)] = \sum_{x=0}^n x(x-1) \binom{n}{x} p^x (1-p)^{n-x} = \sum_{x=2}^n x(x-1) \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \\ &= n(n-1)p^2 \sum_{x=2}^n \frac{(n-2)!}{(x-2)!(n-x)!} p^{x-2} (1-p)^{n-x} = n(n-1)p^2 \end{aligned}$$

$$\begin{aligned} \text{var}(X) &= \mathbf{E}[X^2] - (\mathbf{E}[X])^2 = \mathbf{E}[X^2 - X] + \mathbf{E}[X] - (\mathbf{E}[X])^2 \\ &= n(n-1)p^2 + np - n^2 p^2 = np(1-p) \end{aligned}$$

Mean and Variance of the **Geometric**

- Consider a **geometric** random variable with a PMF

$$\begin{aligned}
 p_X(x) &= (1-p)^{x-1} p, \quad x = 1, 2, \dots, \\
 \mathbf{E}[X] &= \sum_x x p_X(x) = \sum_{x=0}^{\infty} x (1-p)^{x-1} p = p \sum_{x=1}^{\infty} x q^{x-1} \quad (\text{let } q = 1-p < 1) \\
 &= p \frac{d\left(\sum_{x=1}^{\infty} q^x\right)}{dq} = p \frac{d\left(\frac{1}{1-q}\right)}{dq} = p \frac{1}{(1-q)^2} = \frac{1}{p}
 \end{aligned}$$

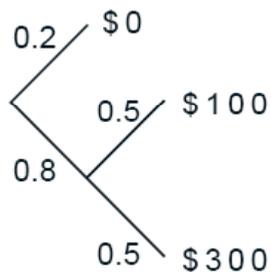
$$\mathbf{E}[X^2] = \mathbf{E}[X^2 - X] + \mathbf{E}[X] \quad (\text{to be verified later on!})$$

$$\begin{aligned}
 \mathbf{E}[X^2 - X] &= \mathbf{E}[X(X-1)] = \sum_{x=0}^{\infty} x(x-1)(1-p)^{x-1} p = pq \sum_{x=2}^{\infty} x(x-1)q^{x-2} \quad (\text{let } q = 1-p < 1) \\
 &= pq \sum_{x=2}^{\infty} x(x-1)q^{x-2} = pq \frac{d^2\left(\frac{1}{1-q}\right)}{d^2q} = pq \frac{2}{(1-q)^3} = \frac{2(1-p)}{p^2}
 \end{aligned}$$

$$\begin{aligned}
 \text{var}(X) &= \mathbf{E}[X^2] - (\mathbf{E}[X])^2 = \mathbf{E}[X^2 - X] + \mathbf{E}[X] - (\mathbf{E}[X])^2 \\
 &= \frac{2(1-p)}{p^2} + \frac{1}{p} - \frac{1}{p^2} = \frac{(1-p)}{p^2}
 \end{aligned}$$

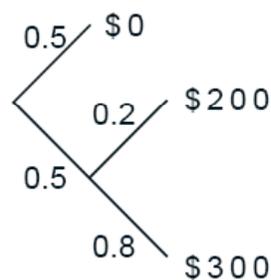
An Example

- **Example 2.3: The Quiz Problem.** Consider a game where a person is given two questions and must decide which question to answer first
 - Question 1 will be answered correctly with probability 0.8, and the person will then receive as prize \$100
 - While question 2 will be answered correctly with probability 0.5, and the person will then receive as prize \$200
 - If the first question attempted is answered incorrectly, the quiz terminates
 - Which question should be answered first to maximize the expected value of the total prize money received?



Question 1
Answered 1st

X



Question 2
Answered 1st

Y

$$P_X(x) = \begin{cases} 0.2 & , x = 0 \\ 0.8 \times 0.5 & , x = 100 \\ 0.8 \times 0.5 & , x = 300 \end{cases} \quad \mathbf{E}[X] = 0.2 \times 0 + 0.4 \times 100 + 0.4 \times 300 = 160$$

$$P_Y(y) = \begin{cases} 0.5 & , y = 0 \\ 0.2 \times 0.5 & , y = 200 \\ 0.8 \times 0.5 & , y = 300 \end{cases} \quad \mathbf{E}[Y] = 0.5 \times 0 + 0.1 \times 200 + 0.4 \times 300 = 140$$

Recitation

- SECTION 2.4 Expectation, Mean, Variance
 - Problems 18, 19, 21, 24