CS70 - Lecture 19 Notes

Name: Felix Su SID: 25794773

Spring 2016 GSI: Gerald Zhang

Random Variables

Real number values assigned to each outcome

- **Definition:** A function X that assigns a real number $X(\omega)$ to each $\omega \in \Omega$
- Set of outcomes s.t. the RV assigned to that outcome $X(\omega)$ is some value $a \in \mathbb{R}$
  - Defined as the inverse image of number (RV) a under the function X.
  - Ex: Two dice roll, RV = total of both dice = 4 (3 possible outcomes) $X^{-1}(4) = \{(1,3),(2,2),(3,1)\}$
- Set of outcomes s.t. the RV to that outcome $X(\omega)$ is some value in $A \in \mathbb{R}$
- Probability that RV $X = a$ is the same as the probability of getting an outcome that maps to a
- The **distribution** of RV X is the set of possible RV values paired with their respective probabilities.

Distribution

All coordinate pairs of X $(\mathrm{RV}, \Pr[\mathrm{RV}])$

Combining Random Variables

- Let X, Y, Z be RVs on Ω and function $g : \mathbb{R}^3 \mapsto \mathbb{R}$
- $g(X, Y, Z)$ = RV that assigns value $g(X(\omega), Y(\omega), X(\omega))$ to ω
- Ex: Three dice roll; X, Y, Z = values of each die; $g(X, Y, Z)$ = max value of X, Y, Z

Set of outcomes s.t. the RV assigned to that outcome $X(\omega)$ is some value $a \in \mathbb{R}$:

$$X^{-1}(a) := \{\omega \in \Omega | X(\omega) = a\} \tag{1}$$

Set of outcomes s.t. the RV to that outcome $X(\omega)$ is some value in $A \in \mathbb{R}$:

$$X^{-1}(A) := \{\omega \in \Omega | X(\omega) = A, A \in \mathbb{R}\} \tag{2}$$

Probability that RV $X = a$ is the same as the probability of getting an outcome that maps to a:

$$\Pr[X = a] = \Pr[X^{-1}(a)] \text{ and } \Pr[X = A] = \Pr[X^{-1}(A)] \tag{3}$$

Distribution

$$\{(a, \Pr[X = a]) : a \in \mathscr{A}\} \text{ where } \mathscr{A} = \{X(\omega), \omega \in \Omega\} \tag{4}$$

Expectation

> 1. Multiply each RV in the distribution of X with its respective probability
> 2. Sum all products

- **Definition:** the **expected value** (meanor expectation) of a random variable X
- Not a common value: Expected value may not be a possible value of X

Law of Large Numbers

> Expectation = average value per experiment if it is performed many times

Expected Value:

$$E[X] = \sum_a a \times \Pr[X = a] \tag{5}$$

Thm: Can sum over outcomes instead of RVs

$$E[X] = \sum_\omega X(\omega) \times \Pr[\omega] \tag{6}$$

Law of Large Numbers: When $n >> 1$

$$E[X] = \frac{X_1 + \cdots + X_n}{n} \tag{7}$$

Indicators

> Random variable that is 1 when ω is in desired event A and 0 otherwise

- **Definition:** Let A = event; **Indicator** of event A = RV X:

$$X = \begin{cases} 1, & \text{if } \omega \in A \\ 0, & \text{if } \omega \notin A \end{cases}$$

Expectation of Indicator

$$E[X] = 1 \times \Pr[X = 1] + 0 \times \Pr[X = 0] = \Pr[A] \tag{8}$$

Alternative form of indicator

$$X(\omega) = 1\{\omega \in A\} \text{ or } 1_A(\omega) \tag{9}$$

$$X = 1_A \tag{10}$$

Linearity of Expectation

- Expectation is linear

Examples

Roll dice n times

- X_m = number of dots on roll m; $X = X_1 + \cdots + X_n$ = total number of dots after n rolls
- $E[X] = E[X_1 + \cdots + X_n]$
- $= E[X_1] + \cdots + E[X_n]$ (by linearity)
- $= nE[X_1]$ because all X_m have the same distribution
- $E[X_1] = 1 \times \frac{1}{6} + \cdots + 6 \times \frac{1}{6} = \frac{6\times 7}{2} \times \frac{1}{6} = \frac{7}{2}$

Flip n coins with heads prob. $= p$ and RV X = no. of heads

- Hard method:
  - $\Pr[X = i] = \binom{n}{i} p^i (1-p)^{n-i}$
  - $E[X] = \sum_i i \times \Pr[X = i] = \sum_i i \times \binom{n}{i} p^i (1-p)^{n-i}$
- Linearity Method:
  - Used X_i as an indicator: 1 if ith flip is heads, 0 otherwise
  - $E[X_i] = 1 \times \Pr[H] + 0 \times \Pr[T] = p$
  - $X = X_1 + \cdots + X_n$
  - $E[X] = E[X_1] + E[X_2] + \cdots + E[X_n] = n \times E[X_i] = np$

Linear Expectation

$$E[a_1 X_1 + \cdots + a_n X_n] = a_1 E[X_1] + \cdots + a_n E[X_n] \tag{11}$$

Union of Indicators

$$1_{A\cup B}(\omega) = 1_A(\omega) + 1_B(\omega) - 1_{A\cap B}(\omega) \tag{12}$$

Probability of Event = Expected value of Indicator RV

$$\Pr[A] = E[1_A] \tag{13}$$

Calculating E[g(x)]

- Let $Y = g(X)$. Assume we know the distribution of X
- Method 1 **(bad)**: Calculate distribution of Y
  - $\Pr[Y = y] = \Pr[X \in g^{-1}(y)]$ where $g^{-1}(x) = \{x \in \mathbb{R} : g(x) = y\}$
- Method 2 **(good)**: Use following Theorem
  - $E[g(X)] = \sum_x g(x) \Pr[X = x]$

Method 2 Example

Let X be uniform in $\{-2, -1, 0, 1, 2, 3\}$; $g(X) = X^2$
$E[g(X)] = \sum_{x=-2}^{3} x^2 \frac{1}{6} = \frac{19}{6}$

Calculate E[g(X)]

$$E[g(X)] = \sum_x g(x)\Pr[X = x] \tag{14}$$

Calculate E[g(X,Y,Z)]

$$E[g(X,Y,Z)] = \sum_{x,y,z} g(X,Y,Z)\Pr[X = x, Y = y, Z = z] \tag{15}$$

Least Squares

> 1. Least Squares: $(X - a)^2$ is used to denote the amount of error
> 2. $a = E[X]$ minimizes $E[(X - a)^2]$, so it is a god guess for X

- **Thm:** The value of a that minimizes $E[(X - a)^2]$ is $a = E[X]$ $\therefore$ if you only know the distribution of X, $E[X]$ is a good guess for X

Least Absolute Deviation

> 1. Least Absolute Deviation: $|X - a|$ is used to denote the amount of error
> 2. $a =$ median of X minimizes $E[|X - a|]$, so it is a god guess for X

- **Thm:** The value of a that minimizes $E[|X - a|]$ is $a =$ median of X $\therefore$ if you only know the distribution of X, the median of X is a good guess for X

Monotonicity

> Mean value of a bigger RV is bigger than the mean value of a smaller RV

- Let X, Y be 2 RVs on Ω
- $X \leq Y$ is $X(\omega)$ is always less than $Y(\omega)$ for all $\omega \in \Omega$, vice verse for $X \geq Y$
- $X \geq a$ for some constant a if $X(\omega)$ is always greater than a
- **Facts**
  - $X \geq 0 \Rightarrow E[X] \geq 0$
  - $X \leq Y \Rightarrow E[X] \leq Y$

Uniform Distribution

> RV X is equally likely to take on any of its values

- X is uniformly distributed in $\{1, 2, \cdots, n\}$ if $\Pr[X = m] = \frac{1}{n}$ for $m = 1, 2, 3 \cdots, n$
- $E[X] = \sum_{m=1}^{n} m\Pr[X = m] = \sum_{m=1}^{n} m \times \frac{1}{n} = \frac{1}{n}\frac{n(n+1)}{2} = \frac{n+1}{2}$

Geometric Distribution

> 1. Flip a coin with $\Pr[H] = p$ until you get H
> 2. Geom. Dist. w/ Parameter p: $\Pr[X = n] = (1 - p)^{n-1}p, n \geq 1$
> 3. Mean value $E[X]$ will increase as p become **smaller** and vice versa

- Flip a coin with $\Pr[H] = p$ until you get H
- Let X = no. of flips until first H
- $X(\omega_n) = n$
- $\Pr[X = n] = (1 - p)^{n-1}p, n \geq 1$

Geometric Distribution w/ Parameter p:

$$\Pr[X = n] = (1 - p)^{n-1}p, n \geq 1 \tag{16}$$

Sum of Geometric Series:

$$\text{if } |a| < 1, S := \sum_{n=0}^{\infty} a^n = \frac{1}{1 - a} \tag{17}$$