Lecture 27

Produced by Dr. Worldwide

Welcome to the 305

Binomial Probability

- Consider tossing a coin with probability of heads equal to p a total of 6 times
- Q: What is the probability that we get exactly 3 heads?
- We could express all possible outcomes of tossing a coin 6 times using a tree diagram that goes on forever, but we all have lives
- Let's consider a few of the outcomes (sequences) where we get exactly 3 heads
- If $\mathrm{A} = \text{Event of Exactly 3 Heads, then } \ \mathrm{A} = \{HHHTTT, TTTHHH, HTHTHT, \cdots\}$
- For each outcome where A occurs, the probability is $p^3(1-p)^3$ because each coin flip is independent
- Q: How many such sequences exist where A occurs?

Binomial Probability

- Number of ways in which we can choose *k* items out of *n* distinct things is called n choose k, denoted by $\binom{n}{k}$, and computed by

$$\binom{n}{k} = \frac{\mathrm{n!}}{(\mathrm{n}-\mathrm{k})!\,\mathrm{k}!}$$

where $\mathrm{n}! = \mathrm{n} \times (\mathrm{n}-1) \times (\mathrm{n}-2) \cdots 3 \times 2 \times 1$ (n factorial)

- The numbers $\binom{n}{0}, \binom{n}{1}, \binom{n}{2}, \cdots, \binom{n}{n-1}, \binom{n}{n}$ are called binomial coefficients, since

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$

Binomial Probability

- From coin example,

  $$\text{P(Exactly 3 Heads)} = \text{P(A)} = \binom{6}{3} p^3 (1-p)^3$$

- Bernouilli process is a repetition of fixed number of independent trials with a binary outcome where the probability of each outcome remains constant

- Each trial/experiment is called a Bernouilli trial

- For a Bernouilli process, the probability of k successes in n trials is

  $$\binom{n}{k} p^n (1-p)^{n-k}$$

- These probabilities build the binomial distribution

- Excel formula is $= BINOM.DIST(n, k, p, FALSE)$

Binomial Probability

- From coin example,

  $$\mathrm{P(Exactly\ 3\ Heads)} = \mathrm{P(A)} = \binom{6}{3} p^3 (1-p)^3$$

- Bernouilli process is a repetition of fixed number of independent trials with a binary outcome where the probability of each outcome remains constant

- Each trial/experiment is called a Bernouilli trial

- For a Bernouilli process, the probability of k successes in n trials is

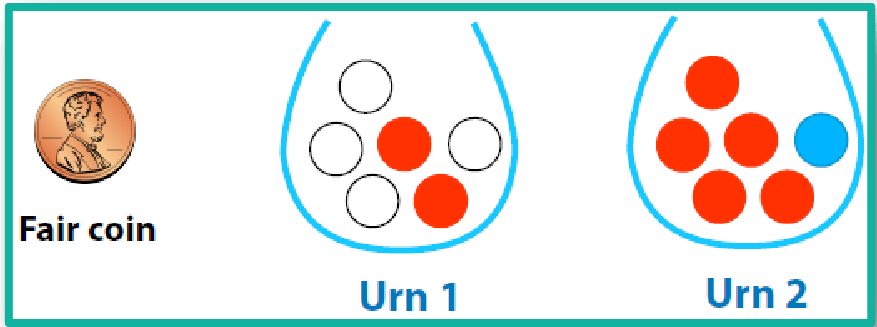
  $$\binom{n}{k} p^n (1-p)^{n-k}$$

- These probabilities build the binomial distribution

- Excel formula is $= BINOM.DIST(n, k, p, FALSE)$

Ex: Balls and Urns

- Suppose we have a fair coin and two urns with six balls each
  - Urn 1 has 4 white balls and 2 red balls
  - Urn 2 has 5 red balls and 1 blue ball

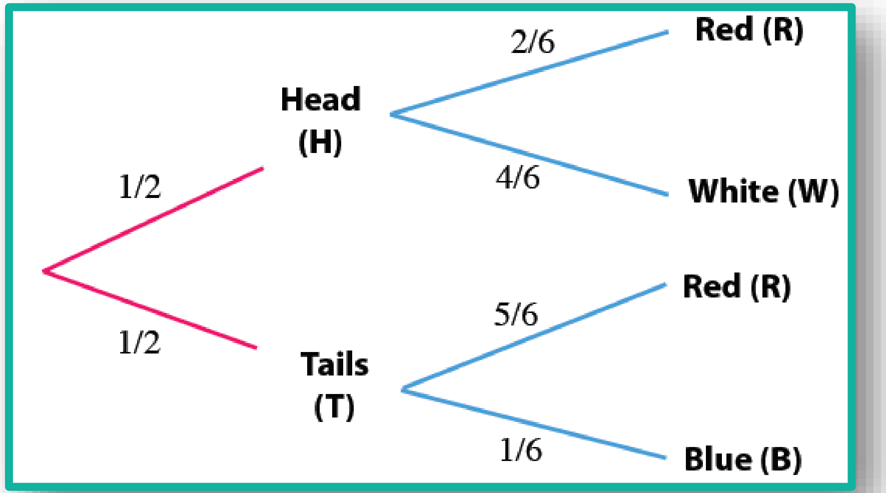


- Experiment
  - Flip a fair coin
  - If heads, draw a ball from urn 1
  - If tails, draw a ball from urn 2

- Q: What are the possible colors of the ball we draw?

Ex: Balls and Urns

- Probability tree to describe all possible outcomes



- List of all outcomes
  - Flip heads and grab red ball ($H \cap R$)
  - Flip heads and grab white ball ($H \cap W$)
  - Flip tails and grab red ball ($T \cap R$)
  - Flip tails and grab blue ball ($T \cap B$)

Ex: Balls and Urns

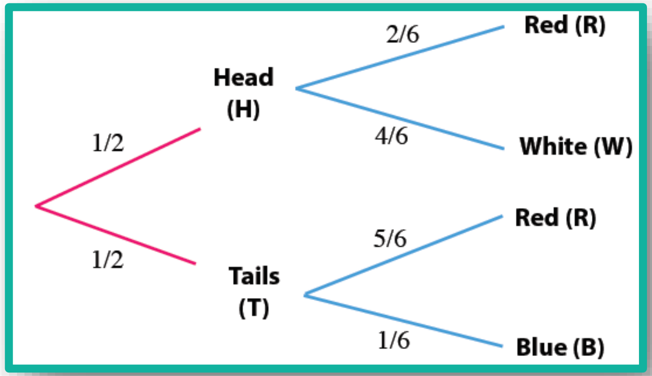
- Probability of all outcomes
  - $P(H \cap R) = \frac{1}{2} \times \frac{2}{6} = \frac{1}{6}$
  - $P(H \cap W) = \frac{1}{2} \times \frac{4}{6} = \frac{1}{3}$
  - $P(T \cap R) = \frac{1}{2} \times \frac{5}{6} = \frac{5}{12}$
  - $P(T \cap B) = \frac{1}{2} \times \frac{1}{6} = \frac{1}{12}$

- Sample space is $\Omega = \{H \cap R, H \cap W, T \cap R, T \cap B\}$ & $\mathrm{P}(\Omega) = \frac{1}{6} + \frac{1}{3} + \frac{5}{12} + \frac{1}{12} = 1$

- Probabilities of each color
  - $P(R) = \frac{1}{6} + \frac{5}{12} = \frac{7}{12}$
  - $P(W) = \frac{1}{3}$
  - $P(B) = \frac{1}{12}$

Ex: Balls and Urns

- Recall the probability tree



- Q: What is probability of a white ball given we flip heads?

  $P(\mathrm{W}|\mathrm{H}) = ??$

- Q: Given we have a white ball, what is the probability the coin flip was heads?

  $P(\mathrm{H}|\mathrm{W}) = ??$

Bayesian Analysis

- Recall the formula for conditional probability

$$P(X|Y) = \frac{P(X \cap Y)}{P(Y)}$$

- Based off this formula,

$$P(Y|X) = \frac{P(X \cap Y)}{P(X)} \quad \rightarrow \quad P(Y|X)P(X) = P(X \cap Y)$$

- With simple substitution,

$$P(X|Y) = \frac{P(Y|X)P(X)}{P(Y)}$$

- Since there are two events, $P(Y) = \mathrm{P}(\mathrm{Y} \cap X) + \mathrm{P}(\mathrm{Y} \cap X^c)$

- This last fact comes from the two branches of a probability tree

Bayesian Analysis

- Based off past statements
  - $\mathrm{P}(\mathrm{Y} \cap X) = \mathrm{P}(\mathrm{Y|X})\mathrm{P}(\mathrm{X})$
  - $\mathrm{P}(\mathrm{Y} \cap X^c) = \mathrm{P}(\mathrm{Y|X}^c)\mathrm{P}(\mathrm{X}^c)$

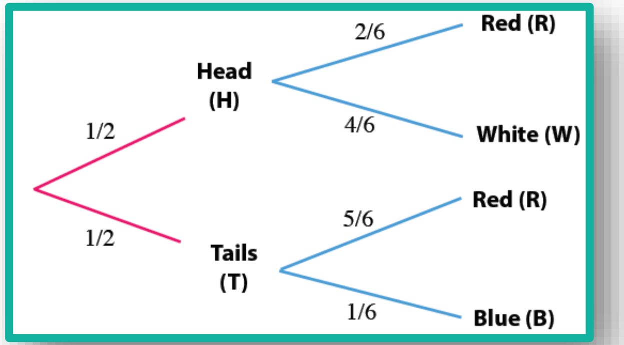
- Bayes' rule

$$P(X|Y) = \frac{P(Y|X)P(X)}{P(Y)} = \frac{P(Y|X)P(X)}{\mathrm{P}(\mathrm{Y|X})\mathrm{P}(\mathrm{X}) + \mathrm{P}(\mathrm{Y|X}^c)\mathrm{P}(\mathrm{X}^c)}$$

- Bayesian analysis
  - Focus on finding posterior probability $P(X|Y)$
  - Prior probability is found in $P(X)$
  - The term $P(Y|X)$ and is typically the likelihood based off gathered data

Ex: Balls and Urns

- Back to the probability tree



- Q: Given we have a white ball, what is the probability the coin flip was heads?

$$P(\mathrm{H|W}) = \frac{\mathrm{P(W|H)P(H)}}{P(W|H)P(H) + P(W|T)P(T)} = \frac{\left({}^{4}/_{6}\right)\left({}^{1}/_{2}\right)}{\left({}^{4}/_{6}\right)\left({}^{1}/_{2}\right) + (0)\left({}^{1}/_{2}\right)} = 1$$

- Q: Given we have a red ball, what is the probability the coin flip was heads?

$$P(\mathrm{H|R}) = \frac{\mathrm{P(R|H)P(H)}}{P(R|H)P(H) + P(R|T)P(T)} = \frac{\left({}^{2}/_{6}\right)\left({}^{1}/_{2}\right)}{\left({}^{2}/_{6}\right)\left({}^{1}/_{2}\right) + \left({}^{5}/_{6}\right)\left({}^{1}/_{2}\right)} = \frac{2}{7}$$

Random Variables

- A random variable X is a function that assigns a real number to each point in the sample space

  $X(s): S \rightarrow \mathbb{R}$

- Examples
  - Tossing a 6-sided die, $X = Result\ of\ Die,\ \ X \in \{1,2,3,4,5,6\}$
  - Observe weather tomorrow, $X = \begin{cases} 1, & if\ it\ rains \\ 0, & if\ it\ doesn't\ rain' \end{cases}\quad X \in \{0,1\}$
  - Wait for the bus, $X = Time\ Spent\ Waiting,\ \ X \in [0, \infty)$
  - Flip a coin 12 times, $X = Number\ of\ Heads,\ \ X \in \{0,1,2,3, \cdots, 12\}$

- Two types
  - A discrete random variable can take on at most a countable number of values
  - A continuous random variable has an uncountable number of values

Random Variables

- Examples of discrete random variables
  - Outcome of a toss of a die (Recoded to binary)
  - Response to a survey question (1 to 5)
  - Number of heads in first 10 tosses
  - Number of cars that pass in front of the Old Well between 8AM and 6PM
  - Grade in school
  - Spread between scores in a basketball game

- Examples of continuous random variables
  - Weight of a baby
  - Height of a giraffe
  - Time a person spends walking per day
  - Age of person

- Q: Can you count the set of integers $\mathbb{Z} = \{0,1,2,3,\cdots\}$?

Discrete Random Variable

- The probability mass function $p(x)$ (pmf) is a function that assigns a probability to every possible value of a discrete random variable X

$$p(x): \{x_1, x_2, \cdots\} \rightarrow [0,1]$$

- Since $p(x)$ is a probability law, $\sum_{i=1}^{\infty} p(x_i) = 1$
- Although $p(x)$ is a function, we interpret it as $p(x) = P(X = x)$
- For any $A \subseteq \{x_1, x_2, \cdots\}$, we have $P(X \in A) = \sum_{x_i \in A} p(x_i)$
- The cumulative distribution function $F(x) = P(X \leq x)$
- For a discrete random variable, $F(x) = P(X \leq x) = \sum_{x_i \leq X} p(x_i)$

Ex: Tossing Coin

- We toss a fair coin 10 times
  - Q: What is the probability there will be at most 2 heads?
  - The random variable $X = number\ of\ heads$
  - The variable X is a binomial random variable

$$p(x) = \binom{10}{x}(0.5)^x(0.5)^{10-x} \text{ where } x \in \{0,1,2,\cdots,10\}$$

  - We want to compute $F(2) = P(X \leq 2)$

$$\begin{aligned} F(2) = P(X \leq 2) &= p(0) + p(1) + p(2) \\ &= (0.5)^{10} + \binom{10}{1}(0.5)^1(0.5)^9 + \binom{10}{2}(0.5)^2(0.5)^8 \\ &= \left(\frac{1}{2}\right)^{10}(1 + 10 + 45) = \frac{7}{2^7} \approx 0.0547 \end{aligned}$$

  - In Excel, $F(2) = BINOM.DIST(2,10,0.5,TRUE)$

$$\begin{aligned} &= BINOM.DIST(0,10,0.5,FALSE) \\ &+ BINOM.DIST(1,10,0.5,FALSE) \\ &+ BINOM.DIST(2,10,0.5,FALSE) \end{aligned}$$

Ex: Rolling Die

- Toss a fair die until the first 6 shows up
  - Q: What is the probability we will need to toss the die at least 4 times?
  - The random variable $X = number\ of\ tosses\ to\ get\ the\ first\ 6$
  - The variable X is a geometric random variable

$$p(x) = \left(\frac{5}{6}\right)^{x-1}\left(\frac{1}{6}\right) \text{ where } x \in \{1,2,3,\cdots\}$$

  - We want to compute $P(X \geq 4) = 1 - P(X < 4) = 1 - F(3)$

$$1 - P(X < 4) = 1 - [p(1) + p(2) + p(3)] = 1 - p(1) - p(2) - p(3)$$

$$= 1 - \frac{1}{6} - \frac{5}{6} \times \frac{1}{6} - \left(\frac{5}{6}\right)^2 \times \frac{1}{6} = \frac{5^3}{6^3} \approx 0.5787$$

  - In Excel, there is not a "GEOM.DIST" function that can be used
  - Can be calculated using Excel's calculator functions

The End

Dale