Miscellaneous Topics in Conditional Probability

Gregory M. Shinault

Goal for this Lecture

Cover a few odds and ends related to conditional probability. Topics covered will include

1. the concept of conditional independence,
2. independence for combinations of independent events, and
3. the Hypergeometric distribution.

The material of this lecture corresponds to Section 2.5 of the textbook.

Conditional Independence

Key Idea

Suppose there is a fair coin, and a coin that has heads on both sides. I randomly choose one of the coins, and flip it twenty times. The result of all twenty tosses is heads.

Do you think the next toss will be heads? Is your answer based on the previous flips? If so, coin flips are supposed to be independent. How do we explain this apparent contradiction? The answer is with conditional independence.

Formal Definition

Let $A_1, A_2, \ldots, A_n$ and B be the events with $\mathbb{P}(B) > 0$. We say that $A_1, A_2, \ldots, A_n$ are **conditionally independent given** B if and only if for any $k \in \{2, \ldots, n\}$ and $1 \leq i_1 < i_2 < \cdots < i_k \leq n$, we have

$$\mathbb{P}(A_{i_1} A_{i_2} \cdots A_{i_k} | B) = \mathbb{P}(A_{i_1}|B)\mathbb{P}(A_{i_2}|B) \cdots \mathbb{P}(A_{i_k}|B).$$

For two events you only check

$$P(A_1 A_2 | B) = P(A_1|B)P(A_2|B).$$

Equivalent Definition

Let $A_1, A_2, \ldots, A_n$ and B be the events with $\mathbb{P}(B) > 0$. We say that $A_1, A_2, \ldots, A_n$ are **conditionally independent given** B if and only if

$$\mathbb{P}(A_1^* A_2^* \cdots A_n^* | B) = \mathbb{P}(A_1^*|B)\mathbb{P}(A_2^*|B) \cdots \mathbb{P}(A_n^*|B).$$

is true for all cases where A_k^* is either A_k or A_k^c.

Example (2.42 from textbook)

Suppose 90% of coins in circulation are fair, and the remaining 10% of coins are biased.

The biased coins give heads with probability $3/5$. We have a random coin and flip it 3 times. What is the proability of getting heads exactly twice?

Example (2.43 from textbook)

We roll two fair dice. Let

$$A = \text{Event that first die is 1 or 2}$$
$$B = \text{Event that 3 appears at least once}$$
$$C = \text{Event that the sum of dice is 5}$$

Show that A and B are dependent, but also A and B are conditionally independent give C.

Independence from Constructed Events

Example (2.41 from textbook)

Let A, B, and C be independent events. Prove that A and $B^c \cup C$ are independent.

The Point of the Previous Example

It might have seemed obvious that A and $B^c \cup C$ are independent, but the actual proof is more complicated than it would seem.

You may use the general fact that constructions like the one in the previous example are independent. However, you should be aware that an elegant proof of this general fact requires mathematical machinery reserved for an advanced course.

The Case for Random Variables

There is a similar fact for random variables: If X, Y, and Z are independent random variables then X and $g(Y, Z)$ are independent for a real-valued function g.

Hypergeometric Distribution

Key Idea

We have introduced the idea of a sample without replacement. This is a common enough tool in statistics that we should identify the formula for its probabilities as one of our special distributions.

Motivation in Terms of Samples

Suppose a population is size N. In this population, there are N_A items of type A and $N_B = N - N_A$ items of type B. We take a sample of size n. Let X denote the number of items in our sample that have type A. Then we say that X has the hypergeometric distribution with parameters N, N_A, n.

- What is the PMF of X is the sample is taken with replacement?
- What is the PMF of X is the sample is taken without replacement?

Definition in Terms of PMF

We say X follows the hypergeometric distribution for with parameters (N, N_A, n) if its probability mass function is

$$p_X(k) = \frac{\binom{N_a}{k}\binom{N-N_A}{n-k}}{\binom{N}{n}}$$

for $k = 0, 1, \ldots, N$.

This is denoted by $X \sim \text{Hypergeom}(N, N_A, n)$.

Example

Suppose X and Y are independent random variables with the distribution $\text{Binomial}(n, p)$. Prove that

$$P(X = k | X + Y = N)$$

is equal to the PMF for the $\text{Hypergeom}(2n, N, n)$ distribution.

Hint: What should the distribution of $X + Y$ be? It is fine to make an appeal to intuition for now. We will formally find the PMF of $X + Y$ later using a technique based on moment generating functions.

The Wrap Up

Summary

- There is a subtle difference between independence and conditional independence. Neither necessarily implies the other.
- If you have mutual independence for several events, it is typically possible to create news events using unions, complements, and intersections that are also independent.
- The **hypergeometric distribution** is the special name we give to the PMF of random variables that count the number of outcomes in a particular category for a sample without replacement.

Next Steps

With the development of conditional probability, we have all of the fundamental probability tools in place. Now we dive into the topic of random variables, and build the remainder of our theory around them.