Applied Statistics

Sofia Triantafillou

sof.triantafillou@gmail.com

University of Crete
Department of Mathematics and Applied Mathematics

March 30, 2022

Lecture Summary

- Discrete Random Variables and Distributions (3.1).
- Expectation (4.1)
- Properties of Expectation (4.2)
- Variance (4.3)
- Joint, Conditional, Marginal Distributions (3.4 – 3.6)
- Conditional Expectation (4.7)

in parentheses: Chapters in the Degroot and Schervish book.

Random Variables

Random Variable

A random variable is a mapping $X : \Omega \to \mathbb{R}$ that assigns a real number $X(\omega)$ to each outcome ω.

Example

Consider the experiment of flipping a coin 10 times, and let $X(\omega)$ denote the number of heads in the outcome ω. For example, if $\omega = HHHTHHTTHT$, $X(\omega) = 6$.

Random Variables

- Why do we need random variables? (easier to work with than original sample space)
- A random variable is NOT a variable (in the algebraic sense).
- A random variable takes a specific value AFTER the experiment is conducted.

Notation

- Letter near the end of the alphabet since it is a variable in the context of the experiment.
- Capital letter to distinguish from algebraic variable.
- Lower case denotes a specific value of the random variable.

Random Variables

- An assignment of a value (number) to every possible outcome.
- Strictly speaking: A function from the sample space Ω to the real numbers.
- It can take discrete or continuous values.
- We will start with discrete values to build our intuition, and then proceed to continuous.

Probability mass function

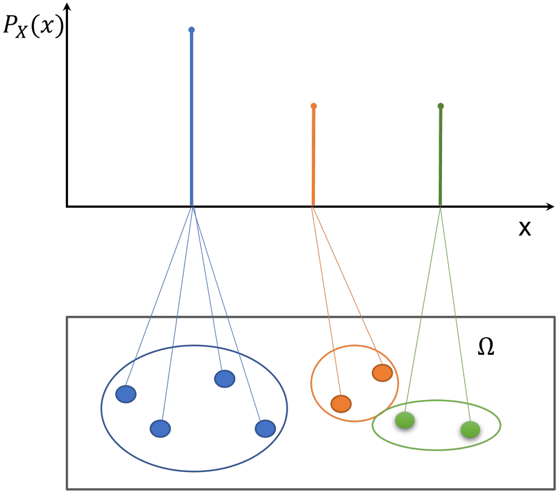
Definition (Probability (mass) function)

If a random variable X has a discrete distribution, the probability function of X is defined as the function p_X such that

$$P_X(x) = \mathbb{P}(X = x) = P(\omega \in \Omega \quad s.t. \quad X(\omega) = x).$$

The closure of the set $x : P_X(x) > 0$ is called the support of the (distribution of) X.

Computing Probability mass functions



Bernoulli distribution

Some distributions come up so often, that they have a name.

Definition

A random variable X that only takes two values 0 (failure) and 1 (success) with $P(X = 1) = p$ has the Bernoulli distribution with parameter p.

Example: You flip a coin one time. X is the result of the coin flip. You can define "Heads" as success and "Tails" as failure (or the other way around, it does not matter) Then X follows a Bernoulli distribution with parameter p.

$$P_X(x) = p^x(1-p)^{1-x}$$

Expectation

Definition

The **expected value** or **mean** or **first moment** of X is defined to be

$$E(X) = \sum_x x P_x(x)$$

assuming that the sum is well-defined.

- We can think of the expectation as the average of a very large number of independent draws from the distribution (IID draws).
- The fact that $E(X) = \sum_{i=1}^n X_i$ is actually a very important theorem we will discuss later.

Examples

- Let $X \sim Bernoulli(p). E(X) = ?$
- Flip a fair coin twice. Let X be the number of heads. $E(X) = ?$

Expectation of a function of a random variable

Sometimes we are interested in the expectation of a function of a random variable $Y = r(X)$. One way to find the expectation of this random variable:

- Find its pmf $P_Y(y)$
- Compute $\sum_y yP_Y(y)$

e.g. Assume X^2 is a discrete random variable with possible values $\{-3, -1, 0, 1, 3\}$ with probabilities $\{\frac{1}{3}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}\}$. Let $Y = X^2$. What is the expectation of Y?

Law Of The Unconscious Statistician

A simple way to compute the expectation of Y in the example above, or any function of random variables, is the law of unconscious statistician (LOTUS).

Theorem

Let $Y = r(X)$. *Then*

$$E(Y) = E(r(x)) = \sum_{\text{all } x} r(x)P_x(x)$$

if the mean exists.

Properties of Expectation

- $E(a) = a$
- $E(aX) = aE(X)$
- $E(aX + b) = aE(X) + b$
- But $E(g(X)) \neq g(E(X))$ in most cases!

Linearity of Expectation

Linearity of expectation is the property that the expected value of the sum of random variables is equal to the sum of their individual expected values, regardless of whether they are independent.

Theorem

Let $X_1, \ldots, X_n$ be a set of random variables. Then

$$E[X_1 + \cdots + X_n] = E[X_1] + \cdots + E[X_n]$$

Prove it for the case of two discrete variables.

Variance of a random variable

Sometimes we are also interested in quantifying how far from the mean

Definition (Variance)

The **variance** of X is defined to be

$$Var(X) = E[(X - E(X))^2]$$

assuming that the sum is well-defined.

Properties of Variance

- $Var(a) = 0$
- $Var(aX) = a^2 Var(X)$
- $Var(aX + b) = a^2 Var(X)$

Examples

- Let $X \sim Bernoulli(p). Var(X) =?$
- Flip a fair coin twice. Let X be the number of heads. $Var(X) =?$

Joint PMFs

y \ x	0	1	2	3
3	1/20	2/20	2/20	
2	2/20	4/20	1/20	2/20
1		2/20	3/20	
0	1/20			

$$P_{X,Y}(x,y) = P(X = x, Y = y) = P(X = x \& Y = y)$$

$$\sum_x \sum y P_{X,Y}(x,y) = 1 \text{ (still a probability mass function)}$$

$$\text{Marginal Probability:} P_X(x) = \sum_y P_{X,Y}(x,y) \text{ (sum over all possible y)}$$

Conditional PMFs

y \ x	0	1	2	3
3	1/20	2/20	2/20	
2	2/20	4/20	1/20	2/20
1		2/20	3/20	
0	1/20			

$y = 2$

$$\text{Conditional Probability:} P_{X|Y}(x|y) = \frac{P_{X,Y}(x,y)}{P_Y(y)}$$

$$\text{e.g., } P_{X|Y}(x|y=2) = \{2/9, 4/9, 1/9, 2/9\}$$

$$\sum_x P_{X|Y}(x|y) = 1 \text{ (still a probability mass function)}$$