Discrete Random Variables

Notes on Bayesian Statistics

Devere Anthony Weaver

---

1 Basic Definition

Definition 1.1 (Discrete Random Variable)**.** *A random variable Y is said to be* ***discrete*** *if it can assume only a finite or countably infinite number of distinct values.*

The collection of probabilities for discrete random variables are useful to know so that we don't have to continuously solve probability problems since many different experiments generate random variables with similar characteristics.

---

2 The Probability Distribution for a Discrete Random Variable

Definition 2.1. *The probability that a random variable Y takes on the particular value y, $P(Y = y)$, is defined as the* ***sum of the probabilities of all sample points in S*** *that are assigned the value of y.*

N.B.

- Observe that the above is often referred to as a **probability function**.

Definition 2.2. *The* ***probability distribution*** *for a discrete random variable Y can be represented by a formula, table, or graph that provides $p(y) = P(Y = y), \forall y$.*

In other words, whatever representation is chosen for the probability function must include the probabilities for each possible value that y can take.

---

3 The Expected Value of a Random Variable or a Function of a Random Variable

Definition 3.1 (Expected Value of a Discrete Random Variable)**.** *Let Y be a discrete random variable with the probability function $p(y)$. The the* ***expected value*** *of Y, $E(Y)$, is defined to be,*

$$E(Y) = \sum_y yp(y). \tag{1}$$

If our $p(y)$ is an accurate characterization of the population frequency distribution, then the expected value is equal to the population mean.

Theorem 3.1. *Let Y be a discrete random variable with probability function $p(y)$ and $g(Y)$ be a real-valued function of Y. Then the expected value of $g(Y)$ is given by*

$$E[g(Y)] = \sum_{\forall y} g(y)p(y). \tag{2}$$

N.B.

- NO proofs of theorem are contained in this document, consult textbook, online, or other resources as needed.

Definition 3.2 (Variance of a Discrete Random Variable)**.** *If Y is a random variable with mean $E(Y) = \mu$, the **variance** of a random variable Y is defined to be the expected value of $(Y - \mu)^2$.*

$$V(Y) = E[(Y-\mu)^2] = \sum_{\forall y} (y-\mu)^2 p(y). \tag{3}$$

*The **standard deviation** of Y is the positive square root of $V(Y)$.*

N.B.

- If $p(y)$ is an accurate characterization of the population frequency distribution, then $V(Y) = \sigma^2$.

Theorem 3.2. *Let Y be a discrete random variable with probability function $p(y)$ and c be a constant. Then $E(c) = c$.*

(i.e. the expected value of a constant is that constant value)

The next theorem states the expected value of a product of a constant and a function of a random variable is equal to the constant times the expected value of the function of the variable.

Theorem 3.3. *Let Y be a discrete random variable with probability function $p(y)$, $g(Y)$ be a function of Y, and c be a constant. Then,*

$$E[cg(Y)] = cE[g(Y)]. \tag{4}$$

The next theorem states the mean or expected value of a sum of functions of a random variable Y is equal to the sum of their respective expected values.

Theorem 3.4. *Let Y be a discrete random variable with probability function $p(y)$ and $g_1(Y), g_2(Y), \dots g_k(Y)$ be k function of Y. Then,*

$$E[g_1(Y) + g_2(Y) + \cdots + g_k(Y)] = E[g_1(Y)] + E[g_2(Y)] + \cdots + E[g_k(Y)]. \tag{5}$$

The three previous theorems can be used to find the variance of a discrete random variable that can reduce the labor in computing it. Of course in reality I'm just going to use a computer...

Theorem 3.5. *Let Y be a discrete random variable with probability function $p(y)$ and mean $E(Y) = \mu$, then*

$$V(Y) = \sigma^2 = E[(Y-\mu)^2] = E(Y^2) - \mu^2. \tag{6}$$

TO DO: Go through the proofs for all the theorems in this section.