

Brian Otieno Odhiambo

**Definition 1.1**

The *mean* of a sample of  $n$  measured responses  $y_1, y_2, \dots, y_n$  is given by:

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$$

The corresponding population mean is denoted  $\mu$ .

**Definition 1.2**

The *variance* of a sample of measurements  $y_1, y_2, \dots, y_n$  is the sum of the square of the differences between the measurements and their mean, divided by  $n - 1$ . Symbolically, the sample variance is:

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2$$

The corresponding population variance is denoted by the symbol  $\sigma^2$ .

**Definition 1.3**

The *standard deviation* of a sample of measurements is the positive square root of the variance; that is,

$$s = \sqrt{s^2}$$

The corresponding population standard deviation is denoted by  $\sigma = \sqrt{\sigma^2}$ .

**Definition 2.6**

Suppose  $S$  is a sample space associated with an experiment. To every event  $A$  in  $S$  ( $A$  is a subset of  $S$ ), we assign a number,  $P(A)$ , called the probability of  $A$ , so that the following axioms hold:

$$\text{Axiom 1: } P(A) \geq 0$$

$$\text{Axiom 2: } P(S) = 1$$

Axiom 3: If  $A_1, A_2, A_3, \dots$  form a sequence of pairwise mutually exclusive events in  $S$  (that is,  $A_i \cap A_j = \emptyset$  if  $i \neq j$ ), then

$$P(S) = P(A_1 \cup A_2 \cup A_3 \cup \dots) = \sum_{i=1}^{\infty} P(A_i)$$

**Definition 2.7**

An ordered arrangement of  $r$  distinct objects is called a *permutation*. The number of ways of ordering  $n$  distinct objects taken  $r$  at a time will be designated by the symbol:

$$P_r^n$$

**Definition 2.8**

The number of *combinations* of  $n$  objects taken  $r$  at a time is the number of subsets, each of size  $r$ , that can be formed from the  $n$  objects. This number will be denoted by:

$$C_r^n \text{ or } \binom{n}{r}$$

#### Theorem 2.4

The number of unordered subsets of size  $r$  chosen (without replacement) from  $n$  available objects is:

$$\binom{n}{r} = C_r^n = \frac{P_r^n}{r!} = \frac{n!}{r!(n-r)!}$$

#### Definition 2.9

The *conditional probability* of an event A, given that an event B has occurred, is equal to:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

provided  $P(B) > 0$ . [The symbol  $P(A|B)$  is read “probability of A given B.”]

#### Definition 2.10

Two events A and B are said to be *independent* if any one of the following holds:

$$P(A|B) = P(A)$$

$$P(B|A) = P(B)$$

$$P(A \cap B) = P(A)P(B)$$

Otherwise, the events are said to be *dependent*.

#### Theorem 2.5

*The Multiplicative Law of Probability* - The probability of the intersection of two events A and B is:

$$\begin{aligned} P(A \cap B) &= P(A)P(B|A) \\ &= P(B)P(A|B) \end{aligned}$$

If A and B are independent, then

$$P(A \cap B) = P(A)P(B)$$

#### Theorem 2.6

*The Additive Law of Probability* - The probability of the union of two events A and B is:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

If A and B are mutually exclusive events,  $P(A \cap B) = 0$  and

$$P(A \cup B) = P(A) + P(B)$$

#### Theorem 2.7

If A is an event, then:

$$P(A) = 1 - P(\overline{A})$$

**Theorem 2.9**

*Bayes' Rule* – Assume that  $\{B_1, B_2, \dots, B_k\}$  is a partition of  $S$  such that  $P(B_i) > 0$ , for  $i = 1, 2, \dots, k$ . Then:

$$P(B_j|A) = \frac{P(A|B_j)P(B_j)}{\sum_{i=1}^k P(A|B_i)P(B_i)}$$

**Definition 3.4**

Let  $Y$  be a discrete random variable with the probability function  $p(y)$ . Then the *expected value* of  $Y$ ,  $E(Y)$ , is defined to be:

$$E(Y) = \sum_y yp(y)$$

**Definition 3.5**

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  $(Y - \mu)^2$ . That is,

$$V(Y) = E[(Y - \mu)^2]$$

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

**Definition 3.7**

A random variable  $Y$  is said to have a *binomial distribution* based on  $n$  trials with success probability  $p$  if and only if:

$$p(y) = \binom{n}{y} p^y q^{n-y}$$

**Theorem 3.7**

Let  $Y$  be a binomial random variable based on  $n$  trials and success probability  $p$ . Then:

$$\mu = E(Y) = np \text{ and } \sigma^2 = V(Y) = npq$$

**Definition 3.8**

A random variable  $Y$  is said to have a *geometric probability distribution* if and only if:

$$p(y) = q^{y-1}p$$

**Theorem 3.8**

If  $Y$  is a random variable with a geometric distribution,

$$\mu = E(Y) = \frac{1}{p} \text{ and } \sigma^2 = V(Y) = \frac{1-p}{p^2}$$

**Definition 3.10**

A random variable  $Y$  is said to have a *hypergeometric probability distribution* if and only if:

$$p(y) = \binom{r}{y} \binom{N-r}{n-y} / \binom{N}{n}$$

where  $y$  is an integer  $0, 1, 2, \dots, n$ , subject to the restrictions  $y \leq r$  and  $n - y \leq N - r$ .

### Definition 3.11

A random variable  $Y$  is said to have a *Poisson probability distribution* if and only if:

$$p(y) = \frac{\lambda^y}{y!} e^{-\lambda}$$

### Theorem 3.11

If  $Y$  is a random variable possessing a Poisson distribution with parameter  $\lambda$ , then:

$$\mu = E(Y) = \lambda \text{ and } \sigma^2 = V(Y) = \lambda$$

### Theorem 3.14

*Tchebysheff's Theorem* - Let  $Y$  be a random variable with mean  $\mu$  and finite variance  $\sigma^2$ . Then, for any constant  $k > 0$ ,

$$P(|Y - \mu| < k\sigma) \geq 1 - \frac{1}{k^2} \text{ or } P(|Y - \mu| \geq k\sigma) \leq \frac{1}{k^2}$$

### Definition 4.1

Let  $Y$  denote any random variable. The *distribution function* of  $Y$ , denoted by  $F(y)$ , is such that

$$F(y) = P(Y \leq y) \text{ for } -\infty < y < \infty$$

### Theorem 4.1

If  $F(y)$  is a distribution function, then

1.  $F(-\infty) \equiv \lim_{y \rightarrow -\infty} F(y) = 0$
2.  $F(\infty) \equiv \lim_{y \rightarrow \infty} F(y) = 1$
3.  $F(y)$  is a nondecreasing function of  $y$ . [If  $y_1$  and  $y_2$  are *any* values such that  $y_1 < y_2$ , then  $F(y_1) \leq F(y_2)$ .]

### Definition 4.2

A random variable  $Y$  with distribution function  $F(y)$  is said to be *continuous* if  $F(y)$  is continuous, for  $-\infty < y < \infty$ .

### Definition 4.3

Let  $F(y)$  be the distribution function for a continuous random variable  $Y$ . Then  $f(y)$ , given by

$$f(y) = \frac{dF(y)}{dy} = F'(y)$$

wherever the derivative exists, is called the *probability density function* for the random variable  $Y$ .

#### **Theorem 4.2**

*Properties of a Density Function* – If  $f(y)$  is a density function for a continuous random variable, then

1.  $f(y) \geq 0$  for all  $y$ ,  $-\infty < y < \infty$
2.  $\int_{-\infty}^{\infty} f(y)dy = 1$

#### **Theorem 4.3**

If the random variable  $Y$  has density function  $f(y)$  and  $a < b$ , then the probability that  $Y$  falls in the interval  $[a, b]$  is

$$P(a \leq Y \leq b) = \int_a^b f(y)dy$$

#### **Definition 4.5**

The expected value of a continuous random variable  $Y$  is

$$E(Y) = \int_{-\infty}^{\infty} yf(y)dy,$$

provided that the integral exists.