

Stats Library

Christopher Schretzmann

April 2024

Definition 1.1 Mean

This finds the mean of a sample:

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

Definition 1.2 Variance

This finds the variance of a sample:

$$s^2 = \frac{1}{n-1} \sum_{i=0}^n (y_i - \bar{y})^2$$

Definition 1.3: Standard Deviation

This finds the standard deviation of the sample:

$$s = \sqrt{\frac{\sum_{i=0}^n (x_i - \bar{x})^2}{n-1}}$$

This sample finds the standard deviation of the population

$$s = \sqrt{\frac{\sum_{i=0}^n (x_i - m)^2}{n}}$$

Theorem 2.1: MxN Rule

With m elements a_1, a_2, \dots, a_m and n elements b_1, b_2, \dots, b_n , it is possible to form $mn = mn$ pairs containing one element from each group.

Definition 2.2: Permutations

This finds the permutations of n distinct objects taken r times:

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

Theorem 2.3 Partition

The number of ways partitioning n distinct objects into k distinct groups where each object appears only once:

$$N = \frac{n!}{n_1!n_2!\dots n_k!}$$

Definition 2.8 Combinations

This finds the number of combinations of n objects taken of r amount:

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

Definition 2.9: Conditional Probability

This finds the conditional probability of event A , given that event B has occurred:

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

Definition 2.10: Independent Probability

Events A and B are independent if any of the following holds:

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

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

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

Otherwise, the events are dependent.

Theorem 2.8: Theorem of Total Probability

Theorem of Total Probability:

$$P(A) = \sum_{i=1}^n P(A|B_i)P(B_i)$$

Assuming $P(B_i) > 0$ for $i = 1, \dots, n$

Theorem 2.9: Baye's Rule

Baye's Rule:

Assume that B_1, B_2, \dots, B_k is a partition of S (see Definition 2.11) such that $P(B_i) > 0$, for $i = 1, 2, \dots, k$. Then

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

Definition 3.4: Discrete Random Variable

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: Expected Value and Standard Deviation of 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 $(Y - \mu)^2$. That is,

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

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

Definition 3.7: Binomial Probability Distribution

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

$$p(y) = \binom{n}{y} p^y q^{n-y}, \quad y = 0, 1, 2, \dots, n \text{ and } 0 \leq p \leq 1.$$

Theorem 3.7: Expected Value and Variance of Binomial Probability Distribution

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

$$\begin{aligned} \text{Expected Value: } \mu &= E(Y) = np \\ \text{Variance: } \sigma^2 &= V(Y) = npq \end{aligned}$$

Definition 3.8: Geometric Probability Distribution

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

$$p(y) = q^{y-1}p, \quad y = 1, 2, 3, \dots, 0 \leq p \leq 1.$$

Theorem 3.8: Expected Value and Variance of Geometric Probability Distribution

If Y is a random variable with a geometric distribution,

$$\begin{aligned} \text{Expected Value: } \mu &= E(Y) = \frac{1}{p} \\ \text{Variance: } \sigma^2 &= V(Y) = \frac{1-p}{p^2} \end{aligned}$$

Definition 3.10: Hypergeometric Probability Distribution

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

$$\frac{\binom{p}{y} \binom{N-r}{n-y}}{\binom{N}{n}}$$

Theorem 3.10: Expected Value and Variance of Hypergeometric Distribution

If Y is a random variable with a hypergeometric distribution,

$$\begin{aligned} \text{Expected Value: } \mu &= E(Y) = \frac{nr}{N} \\ \text{Variance: } \sigma^2 &= V(Y) = n \left(\frac{r}{N} \right) \left(\frac{N-r}{N} \right) \left(\frac{N-n}{N-1} \right) \end{aligned}$$

Definition 3.9: Negative Binomial Probability Distribution

Finds the negative binomial probability distribution

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

Theorem 3.9: Expected Value and Variance of Negative Binomial Probability Distribution

If Y is a random variable with a negative binomial distribution,

$$\begin{aligned}\text{Expected Value: } \mu &= E(Y) = \frac{r}{p} \\ \text{Variance: } \sigma^2 &= V(Y) = \frac{r(1-p)}{p^2}\end{aligned}$$

Definition 3.11: Poisson Probability Distribution

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

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

Where λ = average number of successes in a given interval

Theorem 3.11: Expected Value and Variance of Poisson Probability Distribution

$$\begin{aligned}\text{Expected Value: } \mu &= E(Y) = \lambda \\ \text{Variance: } \sigma^2 &= V(Y) = \lambda\end{aligned}$$

Theorem 3.14: Tchebysheff's Theorem

Let Y be a random variable with mean μ and finite variance σ^2 . Then, for any constant $k > 0$,

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

or

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

Definition 4.1: Distribution Function of Continuous Random Variable

Let Y denote any random variable. The *distribution function* of Y , denoted by $F(y)$, is such that $F(y) = P(Y \leq y)$ for $-\infty < y < \infty$.

Definition 4.3: Probability Density Function of Continuous Random Variable

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 4.3: Probability of Density Function in a Range

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: Expected Value and Variance of Continuous Random Variable

The expected value of a continuous random variable Y is

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

The variance of a continuous random variable Y is

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

Definition 4.6: Continuous Uniform Probability Distribution

If $\theta_1 < \theta_2$, a random variable Y is said to have a continuous uniform probability distribution on the interval (θ_1, θ_2) if and only if the density function of Y is

$$f(y) = \begin{cases} \frac{1}{\theta_2 - \theta_1} & \theta_1 \leq y \leq \theta_2 \\ 0, & \text{elsewhere,} \end{cases}$$

Theorem 4.6: Expected Value and Variance of Uniform Probability Distribution

If $\theta_1 < \theta_2$ and Y is a random variable uniformly distributed on the interval (θ_1, θ_2) , then

$$\text{Expected Value: } \mu = E(Y) = \frac{\theta_1 + \theta_2}{2}$$

$$\text{Variance: } \sigma^2 = V(Y) = \frac{(\theta_2 - \theta_1)^2}{12}$$

Theorem 4.11: Exponential Distribution

The gamma density function in which $\alpha = 1$ is called the exponential density function.

A random variable Y is said to have an exponential distribution with parameter $\beta > 0$ if and only if the density function of Y is

$$f(y) = \begin{cases} \frac{1}{\beta} e^{-\frac{y}{\beta}}, & 0 \leq y < \infty \\ 0, & \text{elsewhere,} \end{cases}$$

Expected Value and Variance of Right Skewed Distribution

$$\text{Expected Value: } \mu = E(Y) = \alpha\beta$$

$$\text{Variance: } \sigma^2 = V(Y) = \alpha\beta^2$$

Theorem 4.10: Expected Value and Variance of Exponential Distribution

If Y is an exponential random variable with parameter β , then

$$\text{Expected Value: } \mu = E(Y) = \beta$$

$$\text{Variance: } \sigma^2 = V(Y) = \beta^2$$

The proof follows directly from Theorem 4.8 with $\alpha = 1$.

Definition 5.1: Joint Probability Function of Discrete Variable

Let Y_1 and Y_2 be discrete random variables. The joint (or bivariate) probability function for Y_1 and Y_2 is given by

$$p(y_1, y_2) = P(Y_1 = y_1, Y_2 = y_2), \quad -\infty < y_1 < \infty, -\infty < y_2 < \infty$$

Definition 5.2: Joint Probability Function of Random Variable

For any random variables Y_1 and Y_2 , the joint (bivariate) distribution function $F(y_1, y_2)$ is

$$F(y_1, y_2) = P(Y_1 \leq y_1, Y_2 \leq y_2), \quad -\infty < y_1 < \infty, -\infty < y_2 < \infty$$

Definition 5.4: Marginal Probability Function

Let Y_1 and Y_2 be jointly discrete random variables with probability function $p(y_1, y_2)$. Then the marginal probability functions of Y_1 and Y_2 , respectively, are given by

$$p_1(y_1) = \sum_{\text{all } y_2} p(y_1, y_2) \quad \text{and} \quad p_2(y_2) = \sum_{\text{all } y_1} p(y_1, y_2).$$

Definition 5.4: Marginal Density Function

Let Y_1 and Y_2 be jointly continuous random variables with joint density function $f(y_1, y_2)$. Then the marginal density functions of Y_1 and Y_2 , respectively, are given by

$$f_1(y_1) = \int_{-\infty}^{\infty} f(y_1, y_2) dy_2 \quad \text{and} \quad f_2(y_2) = \int_{-\infty}^{\infty} f(y_1, y_2) dy_1$$

Definition 5.5: Conditional Discrete Probability Function

If Y_1 and Y_2 are jointly discrete random variables with joint probability function $p(y_1, y_2)$ and marginal probability functions $p_1(y_1)$ and $p_2(y_2)$, respectively, then the conditional discrete probability function of Y_1 given Y_2 is

$$p(y_1|y_2) = P(Y_1 = y_1|Y_2 = y_2) = \frac{P(Y_1 = y_1, Y_2 = y_2)}{P(Y_2 = y_2)} = \frac{p(y_1, y_2)}{p_2(y_2)}$$

Definition 5.6: Conditional Distribution Function

If Y_1 and Y_2 are jointly continuous random variables with joint density function $f(y_1, y_2)$, then the conditional distribution function of Y_1 given $Y_2 = y_2$ is

$$F(y_1|y_2) = P(Y_1 \leq y_1 | Y_2 = y_2)$$

Definition 5.7

Let Y_1 and Y_2 be jointly continuous random variables with joint density $f(Y_1, Y_2)$ and marginal densities $f_1(y_1)$ and $f_2(y_2)$, respectively. For any y_2 such that $f_2(y_2) > 0$, the conditional density of Y_1 given $Y_2 = y_2$ is given by

$$f(y_1|y_2) = \frac{f(y_1, y_2)}{f_2(y_2)}$$

and, for any y_1 such that $f_1(y_1) > 0$, the conditional density of Y_1 given $Y_1 = y_1$ is given by

$$f(y_2|y_1) = \frac{f(y_1, y_2)}{f_1(y_1)}$$

Definition 5.8: Independence of Distribution Function

Let Y_1 have distribution function $F_1(y_1)$, Y_2 have distribution function $F_2(y_2)$, and Y_1 and Y_2 have joint distribution function $F(y_1, y_2)$. Then Y_1 and Y_2 are said to be independent if and only if

$$F(y_1, y_2) = F_1(y_1)F_2(y_2)$$

for every pair of real numbers (y_1, y_2) . If Y_1 and Y_2 are not independent, they are said to be dependent.

Theorem 5.5: Split Function

Let Y_1 and Y_2 have a joint density $f(y_1, y_2)$ that is positive if and only if $a \leq y_1 \leq b$ and $c \leq y_2 \leq d$, for constants a, b, c , and d ; and $f(y_1, y_2) = 0$ otherwise. Then Y_1 and Y_2 are independent random variables if and only if

$$f(y_1, y_2) = g(y_1)h(y_2)$$

where $g(y_1)$ is a nonnegative function of y_1 alone and $h(y_2)$ is a nonnegative function of y_2 alone.