1 Introduction

We consider the properties of the expected value and the variance of a continuous random variable. These quantities are defined just as for discrete random variables and share the same properties.

Expected Value

Definition: Term Let X be a real-valued random variable with density function $f(x)$. The expected value $\mu = E(X)$ is defined by $\mu = E(X) = \int_{-\infty}^{+\infty} x f(x) dx$, provided the integral $\int_{-\infty}^{+\infty} |x| f(x) dx$ is finite.

The reader should compare this definition with the corresponding one for discrete random variables. Intuitively, we can interpret $E(X)$, as we did in the previous sections, as the value that we should expect to obtain if we perform a large number of independent experiments and average the resulting values of X. We can summarize the properties of $E(X)$ as follows.

Theorem If X and Y are real-valued random variables and c is any constant, then

$$E(X+Y) = E(X) + E(Y) \tag{1}$$
$$E(cX) = cE(X). \tag{2}$$

More generally, if $X_1, X_2, \ldots, X_n$ are n real-valued random variables, and $c_1, c_2, \ldots, c_n$ are n constants, then

$$E\left(c_1 X_1 + c_2 X_2 + \cdots + c_n X_n\right) = c_1 E\left(X_1\right) + c_2 E\left(X_2\right) + \cdots + c_n E\left(X_n\right).$$

1.

Example 1. *Let X be uniformly distributed on the interval $[0,1]$. Then $E(X) = \int_0^1 x dx = \frac{1}{2}$ It follows that if we choose a large number N of random numbers from $[0,1]$ and take the average, then we can expect that this average should be close to the expected value of $1/2$.*

2.

Example 2. *Let $Z = (x, y)$ denote a point chosen uniformly and randomly from the unit disk, as in the dart game in Example 2.8, and let $X = \left(x^2 + y^2\right)^{1/2}$ be* *the distance from Z to the center of the disk. The density function of X can easily be shown to equal $f(x) = 2x$, so by the definition of expected value,*

$$\begin{aligned} E(X) &= \int_0^1 x f(x)\, dx \\ &= \int_0^1 x(2x)\, dx \\ &= \frac{2}{3}. \end{aligned}$$

3.

Example 3. *In the example of the couple meeting at the Inn , each person arrives at a time which is uniformly distributed between 5:00 and 6:00 PM. The random variable Z under consideration is the length of time the first person has to wait until the second one arrives. It was shown that*

$$f_Z(z) = 2(1 - z) \quad \textit{for } 0 \le z \le 1.$$

Hence,

$$E(Z) = \int_0^1 z f_Z(z)\, dz = \begin{aligned} &\int_0^1 2z(1-z)\, dz \\ &= \left[z^2 - \frac{2}{3} z^3 \right]_0^1 \\ &= \frac{1}{3}. \end{aligned}$$

Expectation of a Function of a Random Variable

Suppose that X is a real-valued random variable and $\phi(x)$ is a continuous function from $\mathbf{R}$ to $\mathbf{R}$. The following theorem is the continuous analogue of Theorem 1.

Theorem: Expectation of a Continuous Function

If X is a real-valued random variable and $\phi : \mathbf{R} \to \mathbf{R}$ is a continuous real-valued function with domain $[a, b]$, then

$$E(\phi(X)) = \int_{-\infty}^{+\infty} \phi(x) f_X(x)\, dx$$

provided the integral exists.

Theorem: Expectation of the Product of Independent Random Variables

Let X and Y be independent real-valued continuous random variables with finite expected values. Then we have

$$E(XY) = E(X)E(Y)$$

Variance

Definition: Variance

Let X be a real-valued random variable with density function $f(x)$. The variance $\sigma^2 = V(X)$ is defined by

$$\sigma^2 = V(X) = E((X - \mu)^2)$$

Some Properties

Linearity of Variance

If X is a real-valued random variable defined on Ω and c is any constant, then

$$V(cX) = c^2 V(X),$$
$$V(X + c) = V(X).$$

Variance in terms of Expectation

If X is a real-valued random variable with $E(X) = \mu$, then

$$V(X) = E(X^2) - \mu^2.$$

Variance of Sum of Independent Random Variables

If X and Y are independent real-valued random variables on Ω, then

$$V(X + Y) = V(X) + V(Y).$$

Examples

Example 1

If X is uniformly distributed on $[0, 1]$, then, using Theorem 6, we have

$$V(X) = \int_0^1 \left(x - \frac{1}{2}\right)^2 dx = \frac{1}{12}.$$

Example 2

Let X be an exponentially distributed random variable with parameter λ. Then the density function of X is

$$f_X(x) = \lambda e^{-\lambda x}.$$

From the definition of expectation and integration by parts, we have

$$E(X) = \int_0^{\infty} x f_X(x)\, dx = \frac{1}{\lambda}.$$

Similarly, using Theorems 1 and 6, we have

$$V(X) = \frac{1}{\lambda^2}.$$

In this case, both $E(X)$ and $V(X)$ are finite if $\lambda > 0$.

Example 3

Let Z be a standard normal random variable with density function

$$f_Z(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

To calculate the variance of Z, we begin by applying Theorem 6:

$$V(Z) = \int_{-\infty}^{+\infty} x^2 f_Z(x)\, dx - \mu^2.$$

If we write x^2 as $x \cdot x$, and integrate by parts, we obtain

$$\frac{1}{\sqrt{2\pi}} \left(-x e^{-x^2/2} \right) \Big|_{-\infty}^{+\infty} + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-x^2/2}\, dx.$$

The first term can be shown to equal 0, and the second term is just the standard normal density integrated over its domain, so its value is 1. Therefore, the variance of the standard normal density equals 1.

Example 4

Let X and Y be independent random variables, each with mean μ and variance σ^2.

$$\text{Expected value of } S = E(X+Y) = \mu + \mu = 2\mu,$$

$$\text{Variance of } S = V(X+Y) = \sigma^2 + \sigma^2 = 2\sigma^2,$$

The sum of two independent random variables X and Y is denoted as $X+Y$. Let μ_X and μ_Y represent their respective means, and σ_X^2 and σ_Y^2 represent their variances.

The expected value (mean) of the sum is given by:

$$E(X+Y) = E(X) + E(Y) = \mu_X + \mu_Y.$$

Now, let's calculate the variance of the sum:

$$\begin{aligned} V(X+Y) &= \mathrm{Var}(X+Y) \\ &= \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X,Y). \end{aligned}$$

Since X and Y are independent, the covariance term becomes zero:

$$V(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y).$$

Substitute the variances:

$$V(X+Y) = \sigma_X^2 + \sigma_Y^2.$$

This result shows that the variance of the sum of independent random variables is the sum of their individual variances.

Next, let's explore the product of two independent random variables, denoted as XY. The expected value of this product is given by:

$$E(XY) = \mu_X \cdot \mu_Y = \mu_X^2.$$

Now, let's calculate the covariance of X and Y, denoted as $\mathrm{Cov}(X,Y)$:

$$\mathrm{Cov}(X,Y) = E((X-\mu_X)(Y-\mu_Y)).$$

Since X and Y are independent, the joint density function is the product of their individual density functions:

$$f_{XY}(x,y) = f_X(x) \cdot f_Y(y).$$

Express the covariance as an integral:

$$\begin{aligned} \mathrm{Cov}(X,Y) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\mu_X)(y-\mu_Y)f_{XY}(x,y)\,dx\,dy \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\mu_X)(y-\mu_Y)f_X(x)\cdot f_Y(y)\,dx\,dy. \end{aligned}$$

Since X and Y are independent, the joint density factorizes:

$$\begin{aligned} \mathrm{Cov}(X,Y) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\mu_X)(y-\mu_Y)f_X(x)\cdot f_Y(y)\,dx\,dy \\ &= \left(\int_{-\infty}^{\infty}(x-\mu_X)f_X(x)\,dx\right)\cdot\left(\int_{-\infty}^{\infty}(y-\mu_Y)f_Y(y)\,dy\right). \end{aligned}$$

Recognizing that the integrals represent the expected values of $X-\mu_X$ and $Y-\mu_Y$, respectively, we can simplify:

$$\mathrm{Cov}(X,Y) = E(X-\mu_X)\cdot E(Y-\mu_Y).$$

Now, let's use this to express the variance of the sum:

$$\begin{aligned} V(X+Y) &= \text{Var}(X+Y) \\ &= \text{Cov}(X+Y, X+Y). \end{aligned}$$

Since covariance is bilinear, we can expand this expression:

$$V(X+Y) = \text{Cov}(X,X) + \text{Cov}(X,Y) + \text{Cov}(Y,X) + \text{Cov}(Y,Y).$$

Simplifying, we get:

$$V(X+Y) = \text{Var}(X) + 2\text{Cov}(X,Y) + \text{Var}(Y).$$

Now, substitute in the known variances and covariance:

$$V(X+Y) = \sigma_X^2 + 2 \cdot 0 + \sigma_Y^2 = 2\sigma_X^2.$$

So, we have shown that the variance of the sum of independent random variables is the sum of their individual variances.

These examples and properties provide insights into the behavior of random variables and expectations, especially when dealing with independent variables.

Gaussian Distribution:

$$\begin{aligned} \text{Mean } (\mu)&: \quad \text{Represents the center of the distribution.} \\ \text{Variance } (\sigma^2)&: \quad \text{A measure of the spread of the distribution.} \\ \text{Standard Deviation } (\sigma)&: \quad \text{The square root of the variance.} \\ \text{CDF}&: \quad F(x) = \frac{1}{2}\left[1 + \text{erf}\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)\right] \\ \text{MGF}&: \quad M_X(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2} \end{aligned}$$

Francis Galton's Law of Regression:

$$\begin{aligned} \text{Regression Coefficient } (a)&: \quad \text{Indicates the strength and direction of the relationship.} \\ \text{Intercept } (b)&: \quad \text{The value of } Y \text{ when } X \text{ is zero.} \\ \text{Covariance and Correlation}&: \quad \text{Cov}(X,Y) = a \cdot \text{Var}(X) \\ & \quad\ r = \frac{\text{Cov}(X,Y)}{\sqrt{\text{Var}(X) \cdot \text{Var}(Y)}} \end{aligned}$$

Gumbel Distribution:

Location (μ) and Scale (β) Parameters: $E(X) = \mu + 0.57721\beta$

$$\text{Var}(X) = \frac{\pi^2}{6}\beta^2$$

Mode: μ

Reliability Function: $R(x) = e^{-e^{-(x-\mu)/\beta}}$

Fréchet Distribution:

Shape (α), Location (β), and Scale (s) Parameters: $E(X) = \begin{cases} \infty & \text{if } \alpha \leq 1 \\ \beta + \frac{s}{\alpha-1} & \text{if } \alpha > 1 \end{cases}$

Mode: $\beta - \frac{s}{\alpha}$ for $\alpha > 1$

MGF: $M_X(t) = \begin{cases} \infty & \text{if } t < 0 \\ \infty & \text{if } t < 0 \end{cases}$

Pearson's Correlation Coefficient:

Interpretation: A value of$r$ close to 1 or -1 indicates a strong linear relationship, while r

Coefficient of Determination (r^2): $r^2 = \frac{\text{Cov}(X,Y)^2}{\text{Var}(X) \cdot \text{Var}(Y)}$

Linear Regression Equation: $Y = b + aX$

2 the problems of estimating the quantity of collection, sorting and treatment of waste (sampling)

2.1 Heterogeneity of Waste:

Waste streams are often heterogeneous, consisting of various materials in different proportions. This makes it challenging to obtain representative samples.

2.2 Spatial Variation:

Waste characteristics can vary significantly across different locations within a given area. Sampling must consider this spatial variation to ensure accurate representation.

2.3 Temporal Variation:

Waste composition can change over time, influenced by factors such as seasons, holidays, and local events. Accurate estimation requires accounting for temporal variations.

2.4 Sampling Methodology:

Selecting an appropriate sampling methodology is crucial. Improper techniques may lead to biased samples, affecting the reliability of estimates.

2.5 Sample Size:

Determining the appropriate sample size is essential. Too small a sample may not be representative, while an excessively large sample may be impractical or costly.

2.6 Data Collection Frequency:

The frequency of data collection should match the rate of waste generation. Infrequent data collection may lead to outdated information, affecting planning and management decisions.

2.7 Data Accuracy and Precision:

The accuracy and precision of measurement instruments and analytical techniques used in waste characterization impact the reliability of the collected data.

2.8 Public Participation:

Involving the community in waste estimation processes can be challenging. Public cooperation is essential for obtaining accurate data, but it requires effective communication and engagement strategies.

2.9 Data Management and Analysis:

Efficient management and analysis of collected data are critical. Advanced statistical methods may be necessary to process complex waste composition data.

2.10 Regulatory Compliance:

Compliance with regulatory standards for waste estimation is important. Failure to meet these standards can have legal and environmental implications

3 quantification of waste per household and number of population

Quantification of Waste

Household Waste

Data Collection:

Collect data on waste generation from a representative sample of households.

Descriptive Statistics:

Analyze the collected data to identify key statistics like mean, median, and standard deviation of waste generated per household.

Probability Distribution:

Choose an appropriate probability distribution (e.g., Gaussian, Poisson) based on the characteristics of the waste generation data.

Estimation:

Use the chosen distribution to estimate the waste generation per household for the entire population.

Population Waste

Population Size:

Obtain accurate data on the total population size in the targeted area.

Scaling Factor:

Use the waste generation per household estimate and scale it up to the total population:

$$\text{Total Waste} = \text{Waste per Household} \times \text{Total Population}$$

Uncertainty and Variability

Confidence Intervals:

Calculate confidence intervals to account for uncertainty in waste generation estimates.

Sensitivity Analysis:

Perform sensitivity analysis to identify factors that significantly affect waste generation estimates.

Quantification of Waste Using Gaussian Distribution

Household Waste

Data Collection:

Collect data on waste generation from a representative sample of households.

Descriptive Statistics:

Analyze the collected data to identify key statistics like mean (μ) and standard deviation (σ) of waste generated per household.

Probability Distribution:

Assume the waste generated per household follows a Gaussian distribution:

$$P(X) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(X-\mu)^2}{2\sigma^2}}$$

Where:

X is the waste generated per household,
μ is the mean,
σ is the standard deviation.

Estimation:

Use the Gaussian distribution parameters (μ and σ) to estimate the waste generation per household for the entire population.

Population Waste

Population Size:

Obtain accurate data on the total population size in the targeted area.

Scaling Factor:

Use the waste generation per household estimate and scale it up to the total population:

$$\text{Total Waste} = \text{Waste per Household} \times \text{Total Population}$$

Example Plot (Gaussian Distribution)

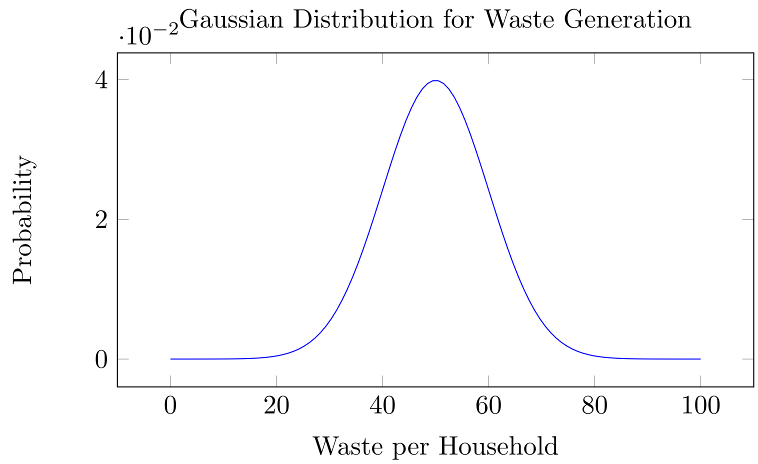

Figure 1: Gaussian distribution example with mean $\mu = 50$ and standard deviation $\sigma = 10$.