Introduction to Statistical Methods in Political Science

Lecture 6: Expected Value and Population Parameters

Ignacio Urbina
Ph.D. Candidate in Political Science

Population Parameters and the Expected Value

Introduction to Population Parameters

- Population parameters (e.g., μ_X and σ_X^2) are numeric, deterministic values that summarize the characteristics of a population's distribution.
- These parameters include measurements of central tendency and spread (i.e., dispersion), providing insights into the general behavior and variability of the population.
- Population parameters are not random variables themselves; they are fixed quantities calculated from the distribution of a random variable.

Expectation: The Expected Value of a Random Variable

- What is the **expected value** or **expectation** of a random variable?
- Simply put, the expected value is the **center** of the distribution.
- But what do we mean by **center**?
  - It is the point that balances the distribution's domain in a precise mathematical sense.
- More formally, the expected value of X is the point where its distribution's **center of mass** lies, balancing the weighted distances to all other points.
- Thus, the **expectation** of X, written as $E[X]$, equals its **population mean**:

$$E[X] = \mu_X.$$

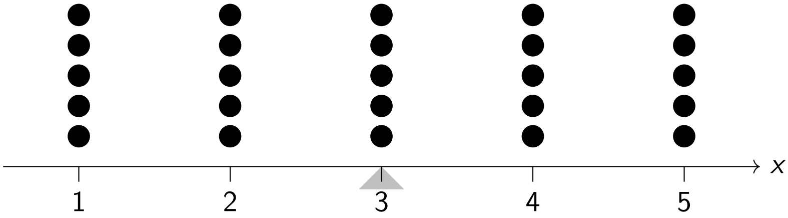
Formal Definition of $E(X)$ - Discrete RV

- Let X be some **discrete random variable (RV)**.
- The *Expected Value* of X is defined as:

$$E(X) = \sum_{x \in D_X} x \cdot p_X(x)$$

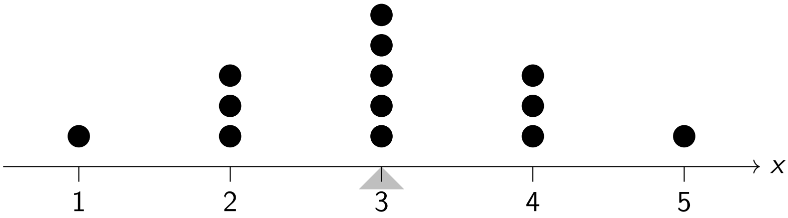
- Here, $p_X(x)$ is the probability mass function (PMF) of X.
- D_X is the domain of X, that is, all the values X can take.
- Note that $E(X)$, at any given point in time, is an unknown constant, not a random variable.

Graphical Example: Symmetric Distribution



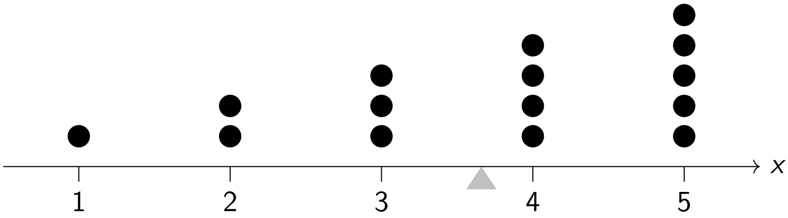
$$\mu_X = 3$$

Graphical Example: Symmetric Distribution



$$\mu_X = 3$$

Graphical Example: Skewed Distribution



$\longrightarrow \mu_X = 3.66$

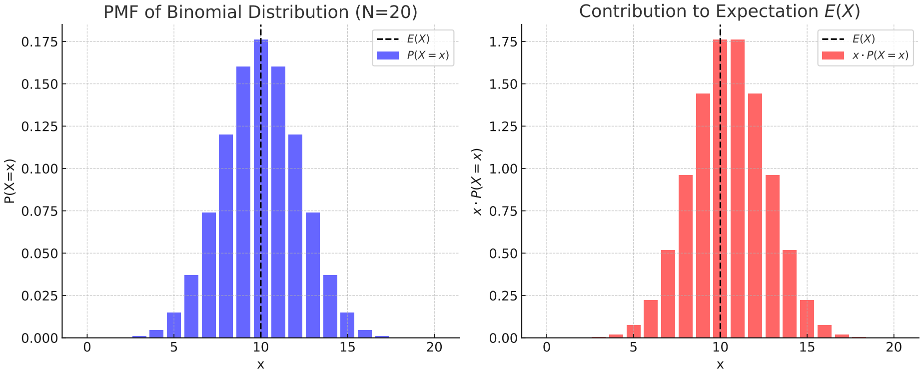
Formal Definition of $E(X)$ - Continous RV

- Let X be a **continuous random variable**.
- The expected value of X is defined as:

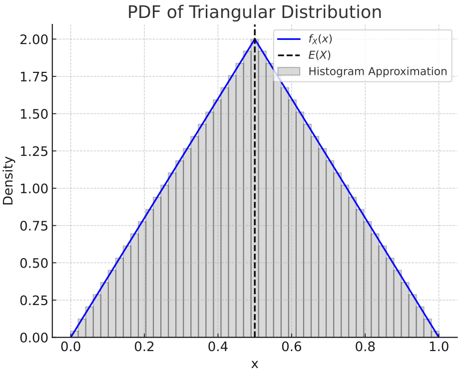
$$E(X) = \int_{-\infty}^{\infty} x \cdot f_X(x)\, dx$$

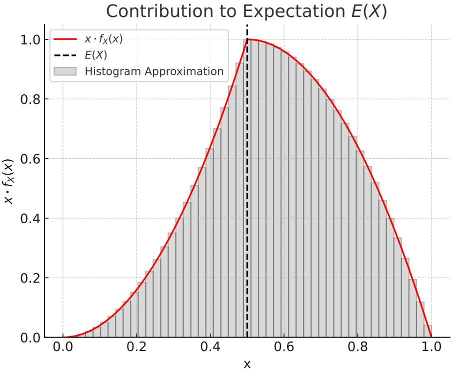
- $f_X(x)$ is the probability density function (PDF) of X.
- The integral captures the total weighted area, just like a sum does in the discrete case.
- Just as a sum adds up discrete contributions, the integral accumulates contributions over a continuous range.

Conceptual Illustration of $E(X)$ - Discrete RV



Conceptual Illustration of $E(X)$ - Continuous RV





Example: Discrete RV Representing Ideology

PMF Table:

x (Ideology)	$P(X = x)$
-5 (Far left)	0.05
-4	0.10
-3	0.15
-2	0.20
-1	0.20
0	0.10
1	0.10
2	0.05
3	0.03
4	0.02
5 (Far right)	0.00

Expectation:

- Definition:

$$E(X) = \sum_{\text{For all } x} x \cdot p_X(x)$$

- Calculation:

$$E(X) = (-5)(0.05) + (-4)(0.10)$$

$$+(-3)(0.15)+(-2)(0.20)+(-1)(0.20)$$

$$+(0)(0.10) + (1)(0.10) + (2)(0.05)$$

$$+(3)(0.03)+(4)(0.02)+(5)(0.00).$$

- Result: $E[X] = -1.33$

Some Remarks on Expectation

- Expectation allows us to define fundamental features of the true (unobserved) distribution of a random variable.
- These features are always **centered** by the weights imposed by the probability distribution of the random variable.
- As statistical analysts, our goal is to use samples to estimate these fundamental features, also called **population parameters**.
- This process forms the foundation of statistical inference, enabling us to draw conclusions about the true distribution from observed data.

Population Variance Expressed Using Expectations

- Analogously to the population mean, we can express the population variance using expectations.
- Population variance measures the **expected quadratic deviation from the population mean.**
- Considering the population probability distribution, we ask: what is the expected quadratic deviation from the mean?
- This approach helps us understand how much the values of a *random variable deviate*, on average, from the mean.

Population Variance: Formal Definition Using Expectation

- Formal Definition:

$$\begin{aligned} V(X) = \sigma_X^2 &= E[(X - E(X))^2] \\ &= E[(X - \mu_X)^2]. \end{aligned}$$

- $V(X)$ and σ_X^2 are alternative notations for "*the variance of X*."
- The population variance is a "centered" measure of the squared deviation of X from its population mean.
  - So, σ_X^2 is also an expected value! (only that we take expectation with regard to $(X - E(X))^2$ and not just X)

Population Variance: Discrete Case

- Variance measures the expected squared deviation from the mean.
- For a discrete random variable:

$$V(X) = E[(X - \mu_X)^2] = \sum_{x \in D_X} (x - \mu_X)^2 \cdot p_X(x)$$

- Where $\mu_X = E(X)$ is the expected value of X.
  - **Note 1**: Population SD is given by $\sigma_X = \sqrt{V(X)} = \sqrt{E[(X - \mu)^2]}$
  - **Note 2**: A general property of expectations is, given some function discrete RV, and any continuous function $g(x)$, then $E[g(x)] = \sum_{x \in D_X} g(x) \cdot p_X(x)$.

Population Variance: Continuous Case

- For a continuous random variable, variance is defined as:

$$V(X) = E[(X - \mu_X)^2] = \int_{-\infty}^{\infty} (x - \mu_X)^2 \cdot f_X(x)\, dx$$

- Where $\mu_X = E(X)$ is the expected value of X, and $f_X(x)$ is the PDF of X.

Population Mean, Variance, and SD for Common RVs Distributions

Distrib.	$\boldsymbol{X}$	$\boldsymbol{E(X)}$	$\boldsymbol{V(X)}$	$\boldsymbol{SD(X) = \sqrt{V(X)}}$
Uniform	$X \sim U[a, b]$	$\frac{a+b}{2}$	$\frac{(b-a+1)^2-1}{12}$	$\sqrt{\frac{(b-a+1)^2-1}{12}}$
Bernoulli	$X \sim Ber(p)$	p	$p(1-p)$	$\sqrt{p(1-p)}$
Binomial	$X \sim Bin(n, p)$	np	$np(1-p)$	$\sqrt{np(1-p)}$
Normal	$X \sim N(\mu, \sigma^2)$	μ	σ^2	σ

Example: Mean and Variance for Uniform Distribution (Die Roll)

PMF for a Die Roll

- $p_X(x) = \frac{1}{6}$ for $x = 1, 2, 3, 4, 5, 6$

Population Mean:

- $E(X) = \mu = \sum_{i=1}^{6}(x_i) \cdot \frac{1}{6} = \frac{1+2+3+4+5+6}{6} = \frac{21}{6} = 3.5.$

Population Variance

- $V(X) = \sigma^2 = E[(X-\mu)^2] =$
  $= \sum_{i=1}^{6}(x_i - \mu)^2 \cdot \frac{1}{6}$
  $= \frac{1}{6}[(1-3.5)^2 + (2-3.5)^2 + (3-3.5)^2 + (4-3.5)^2 + (5-3.5)^2 + (6-3.5)^2]$
  $= \frac{1}{6}[6.25 + 2.25 + 0.25 + 0.25 + 2.25 + 6.25].$
- $\sigma^2 = \frac{1}{6} \times 17.5 = 2.9167$

Population SD

- $SD(X) = \sigma = \sqrt{2.9167} \approx 1.71$

Properties of Expectations and Variances

Let a, b, and c be numeric constants (i.e., not RVs). Let X be any random variable. Then, the following properties hold:

- $E(aX) = aE(X)$
- $E(c) = c$
- $V(X + c) = V(X)$ (Proof)
- $V(bX) = b^2 V(X)$ (Proof)
- $V(X) = E(X^2) - [E(X)]^2$ (Proof)

Linear Combinations (LCs): Definition and Usefulness

Definition:

- A linear combination of random variables X and Y is an expression of the form $aX + bY + c$, where a, b, and c are constants.

Usefulness:

- Helps in deriving properties of combined distributions.
- Essential for understanding linear regression and other statistical models.
- Facilitates transformations to standardize variables.

Motivation: Total Earnings from Two Jobs

- Alex works two part-time jobs:
  - **Job A:** earns a random amount X per week.
  - **Job B:** earns a random amount Y per week.
- Total weekly earnings:

$$Z = X + Y$$

- Key properties:
  - **True Mean Earnings:** $E[Z] = E[X] + E[Y]$
  - **Variance (if independent):** $\mathrm{Var}(Z) = \mathrm{Var}(X) + \mathrm{Var}(Y)$
- If X and Y are correlated, covariance matters!

Properties of Expectations and Variances of LCs

Let a, b, and c be numeric constants (i.e., not RVs). Let X and Y be any two *independent* random variables. Then,

- $E(aX + bY + c) = aE(X) + bE(Y) + c$
- $V(aX + bY + c) = a^2V(X) + b^2V(Y)$ (when X and Y are independent)
- Let $x_1, x_2, \cdots, x_n$ be a series of independent RVs, which could follow different probability distributions. Then,
  - $E[\sum_{i=1}^{n} x_i] = \sum_{i=1}^{n} E[x_i]$
  - $V[\sum_{i=1}^{n} x_i] = \sum_{i=1}^{n} V[x_i]$

Note: We will discuss the case for non-independent X and Y when we cover linear regression.

Practical Example: Z-score Transformation

Assume Y is normally distributed: $Y \sim N(\mu_Y, \sigma_Y^2)$. Let $Z = \frac{Y-\mu}{\sigma}$ be a linear transformation of Y (this is the *Z-score transformation* to Y).

- $E[Z] = E[\frac{Y-\mu}{\sigma}] = E[\frac{Y}{\sigma}] - E[\frac{\mu}{\sigma}] = \frac{E[Y]}{\sigma} - \frac{\mu}{\sigma} = 0$
- $V[Z] = V[\frac{Y-\mu}{\sigma}] = \frac{1}{\sigma^2} V[Y - \mu] = \frac{1}{\sigma^2} \cdot V[Y] = \frac{1}{\sigma^2} \cdot \sigma^2 = 1$
- Therefore, $E[Z] = 0$ and $V[Z] = 1$

Note that this applies to any random variable, regardless of its distribution.

Why use the Z-score Transformation?

Any r.v. X can be expressed as a Z-score, $Z = \frac{X - \mu}{\sigma}$.

- Why would we want to do this? Because units in Z are expressed as deviations from the mean.
- And its unit of measurement (since we divide by the SD) is in SD units.
- Hence, $Z = 2$ means that the original variable's value (e.g., some X) is 2 standard deviations over the mean.

Z-scoring normal random variables

- Teach how to use the Z score and the standard normal to find probabilities of normally distributed RVs

Practical Example: Sample Average as a Linear Combination

- Assume there is a series of independent draws of X, denoted by x_i, all following a normal distribution $X \sim N(\mu, \sigma^2)$.
- Construct a linear combination representing the sample average:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

- Calculate $E[\bar{X}]$:

$$E\left[\bar{X}\right] = E\left[\frac{1}{n} \sum_{i=1}^{n} x_i\right] = \frac{1}{n} E\left[\sum_{i=1}^{n} x_i\right] = \frac{1}{n} \sum_{i=1}^{n} E[x_i] = \frac{1}{n} \sum_{i=1}^{n} \mu = \frac{1}{n} \cdot n\mu = \mu$$

Practical Example: Sample Average as a Linear Combination

- Calculate $V[\bar{X}]$:

$$V\left[\bar{X}\right] = V\left[\frac{1}{n}\sum_{i=1}^{n} x_i\right] = \frac{1}{n^2} V\left[\sum_{i=1}^{n} x_i\right] = \frac{1}{n^2}\sum_{i=1}^{n} V[x_i] = \frac{1}{n^2}\sum_{i=1}^{n}\sigma^2 = \frac{\sigma^2}{n}$$

- Therefore:
  - $E[\bar{X}] = \mu$
  - $V[\bar{X}] = \sigma^2/n$
  - $SD(\bar{X}) = \sqrt{V[\bar{X}]} = \sigma/\sqrt{n}$