Variance and Covariance

Statistics · Module 1 — Probability Foundations

Node 1.6

Position in Knowledge Graph

System:
Statistics

Structural Domain:
Probability Foundations

Node:
1.6

Formal Definition

Let X be a random variable with finite expectation.

Variance

$$\mathrm{Var}(X) = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right].$$

Equivalently,

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2.$$

Covariance

For random variables X, Y:

$$\mathrm{Cov}(X, Y) = \mathbb{E}\left[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])\right].$$

Intuition

Expectation measures center.

Variance measures spread.

Covariance measures joint variation.

If expectation is "balance,"
variance is "dispersion energy."

Covariance tells us whether two variables move together.

Structural Role

Variance and covariance introduce geometry into probability.

They enable:

- Measurement of uncertainty magnitude
- Correlation analysis
- Construction of covariance matrices
- Linear regression
- Multivariate statistics

They transform probability space into an inner-product structure.

Mathematical Development

Key Properties

1. Non-negativity:
$$\mathrm{Var}(X) \geq 0.$$
2. Scaling:
$$\mathrm{Var}(aX) = a^2 \mathrm{Var}(X).$$
3. Variance of Sum:
$$\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X, Y).$$

If X and Y are independent:

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

Correlation

$$\rho(X, Y) = \frac{\mathrm{Cov}(X, Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}.$$

Correlation is dimensionless and bounded:

$$-1 \leq \rho(X, Y) \leq 1.$$

Worked Example

Bernoulli(p)

If $X \sim \mathrm{Bernoulli}(p)$,

$$\mathbb{E}[X] = p,$$

$$\mathrm{Var}(X) = p(1 - p).$$

Maximum variance occurs at $p = 0.5$.

Cross-Links

- **Linear Algebra:** Covariance defines a positive semidefinite matrix.
- **AI:** Covariance matrices appear in PCA and Gaussian models.
- **Economics:** Portfolio variance measures risk.
- **Quantum:** Variance of observables relates to uncertainty principles.

Structural Compression

Invariant:
Variance is always non-negative.

Mechanism:
Quadratic deviation around mean.

Cross-System Echo:
In physics, variance mirrors energy around equilibrium; in ML, variance reflects model sensitivity.

Exercises

1. Prove that variance is non-negative.
2. Show the formula $\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$.
3. Compute variance of a fair die.
4. Prove that if X and Y are independent, then $\mathrm{Cov}(X, Y) = 0$.

Statistics · Module 1 — Probability Foundations · Node 1.6
Concepts: variance-and-covariance, expectation-value

hbar.university — own.your.cognition