VI. Random Variables (RV)

Definition 0.1 (Random Variable)**.** A **Random Variable (RV)** X is a real-valued function defined on the sample space Ω of a probability experiment.

$$X : \Omega \to \mathbb{R}$$

For X to be a valid random variable, the set $\{\omega \in \Omega \mid X(\omega) \leq x\}$ must be an event in the σ-algebra $\mathcal{F}$ for every $x \in \mathbb{R}$. This ensures that we can compute its probability.

Definition 0.2 (Types of Random Variables)**.**

(i) **Discrete Random Variable:** An RV X is discrete if its range (the set of all possible values $X(\omega)$) is finite or countably infinite.

(ii) **Continuous Random Variable:** An RV X is continuous if its range is an uncountably infinite set (typically an interval) and it can be described by a Probability Density Function (PDF).

VII. Discrete Random Variables and Distributions

VII.A. Probability Mass Function (PMF)

Definition 0.3 (Probability Mass Function (PMF))**.** The PMF of a discrete RV X, denoted $p_X(x)$, is defined as:

$$p_X(x) = \mathbb{P}(X = x)$$

Property 0.1 (Properties of a PMF)**.** A function $p(x)$ is a valid PMF for an RV X with range $\{x_1, x_2, \ldots\}$ if:

(i) **Non-negativity:** $p_X(x_i) \geq 0$ for all i.

(ii) **Unit Sum:** $\sum_i p_X(x_i) = \sum_{x \in \text{Range}(X)} p_X(x) = 1$.

VII.B. Cumulative Mass Function (CMF)

Definition 0.4 (Cumulative Mass Function (CMF))**.** The CMF of a discrete RV X, denoted $F_X(x)$, is defined as:

$$F_X(x) = \mathbb{P}(X \leq x) = \sum_{x_i \leq x} p_X(x_i)$$

Property 0.2 (Properties of a CMF)**.**

(i) **Limits:** $\lim_{x \to -\infty} F_X(x) = 0$ and $\lim_{x \to \infty} F_X(x) = 1$.

(ii) **Non-decreasing:** If $a \leq b$, then $F_X(a) \leq F_X(b)$.

(iii) **Right-Continuous:** $F_X(x)$ is a step function that is continuous from the right: $\lim_{\epsilon\to 0^+} F_X(x+\epsilon) = F_X(x)$.

(iv) **PMF from CMF:** $p_X(x) = F_X(x) - F_X(x^-)$, where $F_X(x^-) = \lim_{\epsilon\to 0^+} F_X(x-\epsilon)$ is the limit from the left.

VII.C. Expectation and Variance (Discrete)

Definition 0.5 (Expectation (Mean))**.** The **Expected Value** (or mean) of a discrete RV X is the probability-weighted average of its possible values:

$$\mu = \mathbb{E}[X] = \sum_x x \cdot p_X(x)$$

The expectation exists only if $\sum_x |x| \cdot p_X(x) < \infty$.

Theorem 0.1 (Law of the Unconscious Statistician (LOTUS) - Discrete)**.** Let $g(X)$ be a function of the discrete RV X. The expected value of $g(X)$ is:

$$\mathbb{E}[g(X)] = \sum_x g(x) \cdot p_X(x)$$

Definition 0.6 (Variance)**.** The **Variance** of a discrete RV X, denoted σ^2 or $\mathrm{Var}(X)$, measures the spread of the distribution:

$$\sigma^2 = \mathrm{Var}(X) = \mathbb{E}\left[(X-\mu)^2\right] = \sum_x (x-\mu)^2 \cdot p_X(x)$$

Formula 0.1 (Computational Formula for Variance)**.**

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

The **Standard Deviation** is $\sigma = \sqrt{\mathrm{Var}(X)}$.

Proof (Detailed).[Proof of Computational Formula] $\mathrm{Var}(X) = \mathbb{E}[(X-\mu)^2] = \mathbb{E}[X^2 - 2\mu X + \mu^2]$ By linearity of expectation: $\mathrm{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[2\mu X] + \mathbb{E}[\mu^2]$ Since $\mu = \mathbb{E}[X]$ is a constant: $\mathrm{Var}(X) = \mathbb{E}[X^2] - 2\mu\mathbb{E}[X] + \mu^2 = \mathbb{E}[X^2] - 2\mu(\mu) + \mu^2 = \mathbb{E}[X^2] - \mu^2 = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$. ▲

Property 0.3 (Properties of Expectation and Variance)**.** For constants a and b:

(i) $\mathbb{E}[aX+b] = a\mathbb{E}[X] + b$

(ii) $\mathrm{Var}(aX+b) = a^2\mathrm{Var}(X)$

(iii) $\mathrm{Var}(b) = 0$

VIII. Continuous Random Variables and Distributions

VIII.A. Probability Density Function (PDF)

Definition 0.7 (Probability Density Function (PDF)). An RV X is continuous if there exists a non-negative function $f_X(x)$, the PDF, such that for any set $B \subset \mathbb{R}$:

$$\mathbb{P}\left(X \in B\right) = \int_B f_X(x)dx$$

For an interval $[a, b]$, this means $\mathbb{P}\left(a \leq X \leq b\right) = \int_a^b f_X(x)dx$.

Property 0.4 (Properties of a PDF). A function $f(x)$ is a valid PDF if:

(i) **Non-negativity:** $f_X(x) \geq 0$ for all x.

(ii) **Unit Integral:** $\int_{-\infty}^{\infty} f_X(x)dx = 1$.

Note: For any continuous RV, $\mathbb{P}\left(X = a\right) = \int_a^a f_X(x)dx = 0$.

VIII.B. Cumulative Distribution Function (CDF)

Definition 0.8 (Cumulative Distribution Function (CDF)). The CDF of a continuous RV X, denoted $F_X(x)$, is defined as:

$$F_X(x) = \mathbb{P}\left(X \leq x\right) = \int_{-\infty}^{x} f_X(t)dt$$

Property 0.5 (Properties of a CDF).

(i) **Limits:** $\lim_{x\to-\infty} F_X(x) = 0$ and $\lim_{x\to\infty} F_X(x) = 1$.

(ii) **Non-decreasing:** If $a \leq b$, then $F_X(a) \leq F_X(b)$.

(iii) **Continuous:** $F_X(x)$ is a continuous function.

(iv) **Relationship to PDF:** $f_X(x) = \frac{d}{dx}F_X(x) = F'_X(x)$ at all points where F_X is differentiable.

(v) **Probability from CDF:** $\mathbb{P}\left(a < X \leq b\right) = F_X(b) - F_X(a)$.

VIII.C. Expectation and Variance (Continuous)

Definition 0.9 (Expectation (Mean)). The **Expected Value** of a continuous RV X is:

$$\mu = \mathbb{E}\left[X\right] = \int_{-\infty}^{\infty} x \cdot f_X(x)dx$$

The expectation exists only if $\int_{-\infty}^{\infty} |x| \cdot f_X(x)dx < \infty$.

Theorem 0.2 (Law of the Unconscious Statistician (LOTUS) - Continuous). Let $g(X)$ be a function of the continuous RV X. The expected value of $g(X)$ is:

$$\mathbb{E}\left[g(X)\right] = \int_{-\infty}^{\infty} g(x) \cdot f_X(x)dx$$

Definition 0.10 (Variance)**.** The **Variance** of a continuous RV X is:

$$\sigma^2 = \mathrm{Var}\left(X\right) = \mathbb{E}\left[(X-\mu)^2\right] = \int_{-\infty}^{\infty} (x-\mu)^2 \cdot f_X(x)dx$$

Formula 0.2 (Computational Formula for Variance)**.**

$$\mathrm{Var}\left(X\right) = \mathbb{E}\left[X^2\right] - (\mathbb{E}\left[X\right])^2 = \left(\int_{-\infty}^{\infty} x^2 f_X(x)dx\right) - \left(\int_{-\infty}^{\infty} x f_X(x)dx\right)^2$$

IX. Jointly Distributed Random Variables

IX.A. Joint Distributions (Discrete and Continuous)

Definition 0.11 (Joint PMF (Discrete))**.** For two discrete RVs X and Y, the Joint PMF is:

$$p(x,y) = \mathbb{P}\left(X = x, Y = y\right)$$

Properties: (i) $p(x,y) \geq 0$, (ii) $\sum_x \sum_y p(x,y) = 1$. This is often represented in a **contingency table**.

Definition 0.12 (Joint PDF (Continuous))**.** For two continuous RVs X and Y, the Joint PDF $f(x,y)$ satisfies:

$$\mathbb{P}\left((X,Y) \in A\right) = \iint_A f(x,y)dxdy$$

Properties: (i) $f(x,y) \geq 0$, (ii) $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y)dxdy = 1$.

Definition 0.13 (Joint CDF)**.** $F(x,y) = \mathbb{P}\left(X \leq x, Y \leq y\right)$.

- **Discrete:** $F(x,y) = \sum_{s \leq x} \sum_{t \leq y} p(s,t)$.
- **Continuous:** $F(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(s,t)dtds$.
- **PDF from CDF:** $f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y}$.

IX.B. Marginal Distributions

Definition 0.14 (Marginal Distributions)**.** The distribution of a single RV from a joint distribution.

- **Discrete (Marginal PMF):** $p_X(x) = \sum_y p(x,y) = \mathbb{P}\left(X = x\right)$.
- **Continuous (Marginal PDF):** $f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy$.

IX.C. Independence and Covariance

Definition 0.15 (Statistical Independence of RVs)**.** Two RVs X and Y are **independent** if for all x, y:

(i) **General:** $F(x, y) = F_X(x)F_Y(y)$.

(ii) **Discrete:** $p(x, y) = p_X(x)p_Y(y)$.

(iii) **Continuous:** $f(x, y) = f_X(x)f_Y(y)$.

Theorem 0.3 (Expectation of a Product (Multiplication Theorem))**.** For any two RVs X, Y, and functions g, h:

$$\mathbb{E}[g(X)h(Y)] = \mathbb{E}[g(X)]\,\mathbb{E}[h(Y)] \quad \text{if } X, Y \text{ are independent.}$$

A special case is $\mathbb{E}[XY] = \mathbb{E}[X]\,\mathbb{E}[Y]$.

Definition 0.16 (Covariance)**.** The **Covariance** of X and Y measures their linear relationship:

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

Formula 0.3 (Computational Formula for Covariance)**.**

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

Proof (Detailed).[Proof of Computational Formula] $\mathrm{Cov}(X, Y) = \mathbb{E}[XY - X\mu_Y - Y\mu_X + \mu_X\mu_Y]$ By linearity of expectation: $= \mathbb{E}[XY] - \mathbb{E}[X\mu_Y] - \mathbb{E}[Y\mu_X] + \mathbb{E}[\mu_X\mu_Y] = \mathbb{E}[XY] - \mu_Y\mathbb{E}[X] - \mu_X\mathbb{E}[Y] + \mu_X\mu_Y = \mathbb{E}[XY] - \mu_Y\mu_X - \mu_X\mu_Y + \mu_X\mu_Y = \mathbb{E}[XY] - \mu_X\mu_Y$. ▲

Property 0.6 (Covariance and Independence)**.** If X and Y are independent, then $\mathbb{E}[XY] = \mathbb{E}[X]\,\mathbb{E}[Y]$, so:

$$\mathrm{Cov}(X, Y) = \mathbb{E}[X]\,\mathbb{E}[Y] - \mathbb{E}[X]\,\mathbb{E}[Y] = 0$$

Caution: The converse is NOT true. $\mathrm{Cov}(X, Y) = 0$ (uncorrelated) does not imply independence.

Property 0.7 (Properties of Covariance)**.** (i) $\mathrm{Cov}(X, X) = \mathrm{Var}(X)$

(ii) $\mathrm{Cov}(X, Y) = \mathrm{Cov}(Y, X)$

(iii) $\mathrm{Cov}(aX + b, cY + d) = ac\mathrm{Cov}(X, Y)$

(iv) $\mathrm{Cov}(X + Y, Z) = \mathrm{Cov}(X, Z) + \mathrm{Cov}(Y, Z)$

Formula 0.4 (Variance of a Sum (General))**.**

$$\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X, Y)$$

If X, Y are independent, $\mathrm{Cov}(X, Y) = 0$, so $\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$.

Formula 0.5 (Variance of a Sum (n variables))**.**

$$\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{Var}\left(X_i\right) + 2\sum_{i<j} \mathrm{Cov}\left(X_i, X_j\right)$$

X. Conditional Distributions, Mean, and Variance

X.A. Conditional Distributions

Definition 0.17 (Conditional PMF (Discrete))**.** The conditional PMF of X given $Y = y$ is:

$$p_{X|Y}(x|y) = \mathbb{P}\left(X = x \mid Y = y\right) = \frac{\mathbb{P}\left(X = x, Y = y\right)}{\mathbb{P}\left(Y = y\right)} = \frac{p(x,y)}{p_Y(y)}$$

This is a valid PMF, so $\sum_x p_{X|Y}(x|y) = 1$ for any fixed y.

Definition 0.18 (Conditional PDF (Continuous))**.** The conditional PDF of X given $Y = y$ is:

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)} \quad (\text{where } f_Y(y) > 0)$$

This is a valid PDF, so $\int_{-\infty}^{\infty} f_{X|Y}(x|y)dx = 1$ for any fixed y.

Property 0.8 (Independence and Conditionals)**.** X and Y are independent if and only if $p_{X|Y}(x|y) = p_X(x)$ (or $f_{X|Y}(x|y) = f_X(x)$) for all x, y.

X.B. Conditional Expectation

Definition 0.19 (Conditional Expectation)**.** The **Conditional Expectation** of X given $Y = y$ is the mean of the conditional distribution:

- **Discrete:** $\mathbb{E}\left[X \mid Y = y\right] = \sum_x x \cdot p_{X|Y}(x|y)$
- **Continuous:** $\mathbb{E}\left[X \mid Y = y\right] = \int_{-\infty}^{\infty} x \cdot f_{X|Y}(x|y)dx$

Note: $\mathbb{E}\left[X \mid Y = y\right]$ is a number. The function $g(y) = \mathbb{E}\left[X \mid Y = y\right]$ defines a new random variable $\mathbb{E}\left[X \mid Y\right] = g(Y)$.

Theorem 0.4 (Law of Total Expectation (Adam's Law / Tower Property))**.** The expected value of X is the expected value of its conditional expectation given Y.

$$\mathbb{E}\left[X\right] = \mathbb{E}\left[\mathbb{E}\left[X \mid Y\right]\right]$$

Proof (Detailed).[Proof of Total Expectation (Continuous Case)] The outer expectation is taken with respect to Y.

$$\mathbb{E}\left[\mathbb{E}\left[X \mid Y\right]\right] = \int_{-\infty}^{\infty} \mathbb{E}\left[X \mid Y = y\right] \cdot f_Y(y)dy$$

Substitute the definition of conditional expectation:

$$= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} x \cdot f_{X|Y}(x|y)dx\right] f_Y(y)dy$$

Substitute $f_{X|Y}(x|y) = f(x,y)/f_Y(y)$:

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x \cdot \frac{f(x,y)}{f_Y(y)} f_Y(y)dxdy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x \cdot f(x,y)dydx$$

Reverse the order of integration:

$$= \int_{-\infty}^{\infty} x \left[\int_{-\infty}^{\infty} f(x,y)dy\right] dx$$

The inner integral is the marginal PDF $f_X(x)$:

$$= \int_{-\infty}^{\infty} x \cdot f_X(x)dx = \mathbb{E}\left[X\right]$$

▲

X.C. Conditional Variance

Definition 0.20 (Conditional Variance)**.** The **Conditional Variance** of X given $Y = y$ is the variance of the conditional distribution:

$$\operatorname{Var}\left(X \mid Y = y\right) = \mathbb{E}\left[(X - \mathbb{E}\left[X \mid Y = y\right])^2 \mid Y = y\right]$$

Computational Formula: $\operatorname{Var}\left(X \mid Y = y\right) = \mathbb{E}\left[X^2 \mid Y = y\right] - (\mathbb{E}\left[X \mid Y = y\right])^2$

Theorem 0.5 (Law of Total Variance (Eve's Law))**.** The variance of X is the sum of the expected conditional variance and the variance of the conditional expectation.

$$\operatorname{Var}\left(X\right) = \mathbb{E}\left[\operatorname{Var}\left(X \mid Y\right)\right] + \operatorname{Var}\left(\mathbb{E}\left[X \mid Y\right]\right)$$

Proof (Detailed).[Proof of Total Variance] We prove this by showing the two terms on the RHS sum to $\mathbb{E}\left[X^2\right] - (\mathbb{E}\left[X\right])^2$.

(i) **First Term (Mean of Conditional Variance):** $\mathbb{E}\left[\operatorname{Var}\left(X \mid Y\right)\right] = \mathbb{E}\left[\mathbb{E}\left[X^2 \mid Y\right] - (\mathbb{E}\left[X \mid Y\right])^2\right]$ By linearity of expectation: $\mathbb{E}\left[\operatorname{Var}\left(X \mid Y\right)\right] = \mathbb{E}\left[\mathbb{E}\left[X^2 \mid Y\right]\right] - \mathbb{E}\left[(\mathbb{E}\left[X \mid Y\right])^2\right]$ By the Law of Total Expectation, $\mathbb{E}\left[\mathbb{E}\left[X^2 \mid Y\right]\right] = \mathbb{E}\left[X^2\right]$. So, $\mathbb{E}\left[\operatorname{Var}\left(X \mid Y\right)\right] = \mathbb{E}\left[X^2\right] - \mathbb{E}\left[(\mathbb{E}\left[X \mid Y\right])^2\right]$.

(ii) **Second Term (Variance of Conditional Mean):** Let $g(Y) = \mathbb{E}\left[X \mid Y\right]$. By definition of variance: $\operatorname{Var}\left(\mathbb{E}\left[X \mid Y\right]\right) = \operatorname{Var}\left(g(Y)\right) = \mathbb{E}\left[(g(Y))^2\right] - (\mathbb{E}\left[g(Y)\right])^2$ $\operatorname{Var}\left(\mathbb{E}\left[X \mid Y\right]\right) = \mathbb{E}\left[(\mathbb{E}\left[X \mid Y\right])^2\right] - (\mathbb{E}\left[\mathbb{E}\left[X \mid Y\right]\right])^2$ By the Law of Total Expectation, $\mathbb{E}\left[\mathbb{E}\left[X \mid Y\right]\right] = \mathbb{E}\left[X\right]$. So, $\operatorname{Var}\left(\mathbb{E}\left[X \mid Y\right]\right) = \mathbb{E}\left[(\mathbb{E}\left[X \mid Y\right])^2\right] - (\mathbb{E}\left[X\right])^2$.

(iii) **Summing (i) and (ii):** $(\mathbb{E}[X^2] - \mathbb{E}[(\mathbb{E}[X \mid Y])^2]) + (\mathbb{E}[(\mathbb{E}[X \mid Y])^2] - (\mathbb{E}[X])^2)$ The middle terms cancel, leaving: $\mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \mathrm{Var}(X)$.

▲