

ECE 30200 - Probabilistic Methods in Electrical and Computer Engineering

Zeke Ulrich

June 26, 2025

Contents

<i>Background</i>	<i>2</i>
<i>Series</i>	<i>2</i>
<i>Combinatorics</i>	<i>2</i>
<i>Approximations</i>	<i>2</i>
<i>Calculus</i>	<i>3</i>
<i>Linear Algebra</i>	<i>3</i>
<i>Probability Laws</i>	<i>4</i>
<i>Probability Properties</i>	<i>4</i>
<i>Formal Definitions</i>	<i>5</i>
<i>Outcomes</i>	<i>5</i>
<i>Events</i>	<i>5</i>
<i>Probability Laws</i>	<i>5</i>
<i>Probability Space</i>	<i>5</i>
<i>Probability Properties</i>	<i>6</i>
<i>Random Variables</i>	<i>7</i>
<i>Reference</i>	<i>12</i>
<i>Series</i>	<i>12</i>
<i>Combinatorics</i>	<i>12</i>
<i>Approximations</i>	<i>12</i>
<i>Calculus</i>	<i>13</i>
<i>Linear Algebra</i>	<i>13</i>

Background

The following formulas will be instrumental and may be familiar.

Series

$$\sum_{k=0}^n r^k = \frac{1 - r^{n+1}}{1 - r} \quad (1)$$

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6} \quad (2)$$

$$\sum_{k=1}^{\infty} k r^{k-1} = \frac{1}{(1 - r)^2} \quad (3)$$

Combinatorics

$$\binom{n}{k} = \frac{n!}{k!(n - k)!} \quad (4)$$

$$(a + b)^n = \sum_{k=0}^n \binom{n}{k} a^{n-k} b^k \quad (5)$$

$$\binom{n}{k} + \binom{n}{k-1} = \binom{n+1}{k} \quad (6)$$

$$P(n, k) = \frac{n!}{(n - k)!} \quad (7)$$

where $P(n, k)$ is the number of ways to arrange k objects out of n (permutations).

$$C(n, k) = \binom{n}{k} = \frac{n!}{k!(n - k)!} \quad (8)$$

where $C(n, k)$ is the number of ways to choose k objects out of n (combinations).

Approximations

$$f(x) = f(a) + f'(a)(x - a) + \frac{f''(a)}{2!}(x - a)^2 + \dots \quad (9)$$

$$= \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!} (x - a)^n \quad (10)$$

$$1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots = \sum_{k=0}^{\infty} \frac{x^k}{k!} \quad (11)$$

$$= e^x \quad (12)$$

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots \quad (13)$$

$$= \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} \quad (14)$$

$$\cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots \quad (15)$$

$$= \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} \quad (16)$$

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots \quad (17)$$

$$= \sum_{n=1}^{\infty} (-1)^{n+1} \frac{x^n}{n} \quad (18)$$

Calculus

$$\frac{d}{dx} \int_a^x f(t) dt = f(x) \quad (19)$$

$$\int_a^b f'(x) dx = f(b) - f(a) \quad (20)$$

$$\int f(g(x))g'(x) dx = \int f(u) du \quad (21)$$

$$\int u dv = uv - \int v du \quad (22)$$

$$\int \frac{1}{(x-a)(x-b)} dx = \frac{1}{b-a} \ln \left| \frac{x-a}{x-b} \right| + C \quad (23)$$

Linear Algebra

$$\vec{y} = \beta_1 \vec{x}_1 + \beta_2 \vec{x}_2 + \dots + \beta_N \vec{x}_N \quad (24)$$

$$\langle \vec{a}, \vec{b} \rangle = \vec{a} \vec{b}^T \quad (25)$$

$$= \sum_{i=1}^n a_i b_i \quad (26)$$

where $\langle \vec{a}, \vec{b} \rangle$ denotes the inner product of vectors \vec{a} and \vec{b} .

$$\|\vec{x}\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{1/p} \quad (27)$$

where $\|\vec{x}\|_p$ is the p -norm (or ℓ_p -norm) of vector \vec{x} .

$$\cos(\theta) = \frac{\langle \vec{a}, \vec{b} \rangle}{\|\vec{a}\|_2 \|\vec{b}\|_2} \quad (28)$$

where θ is the angle between vectors \vec{a} and \vec{b} .

$$\hat{\beta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \vec{y} \quad (29)$$

where $\hat{\beta}$ is the vector of least squares coefficients, \mathbf{X} is the data matrix, and \vec{y} is the target vector

Probability Laws

A probability law must satisfy three axioms:

1. Non-negativity: $P(A) \geq 0 \forall A \in F$
2. Normalization: $P(\Omega) = 1$
3. Additivity: For any disjoint subsets $\{A_1, A_2, \dots\}$, it holds that

$$P \left[\bigcup_{n=1}^{\infty} A_n \right] = \sum_{n=1}^{\infty} P[A_n]$$

Probability Properties

$$P[A \cup B] = P[A] + P[B] - P[A \cap B] \quad (30)$$

$$P[A \cup B] \leq P[A] + P[B] \quad (31)$$

$$A \subseteq B \implies P[A] \leq P[B] \quad (32)$$

Formal Definitions

Outcomes

An *outcome* is the result of some *experiment*. If that experiment is flipping a coin, the outcome is either heads or tails. We could express the outcome of heads as H , and the outcome of tails as T . The set of all possible outcomes for an experiment is known as a sample space and is denoted by Ω . In this case $\Omega = \{H, T\}$.

Events

An *event* F is a subset of the sample space Ω . The formal definitions of probability are expressed with set notation. So the event where we have neither heads nor tails is written as $\{\}$. The event of heads could be expressed as $\{H\}$, and the event of tails could be expressed as $\{T\}$. The event of either heads or tails is $\{H, T\}$.

Probability Laws

A *probability law* is a function P that maps an event A to a real number in $[0, 1]$. For the coin example, the probability law might be $P(\{\}) = 0$, $P(\{H\}) = 0.5$, $P(\{T\}) = 0.5$, and $P(\{\Omega\}) = 1$. A probability law must satisfy three axioms:

1. Non-negativity: $P(A) \geq 0 \forall A \in F$
2. Normalization: $P(\Omega) = 1$
3. Additivity: For any disjoint subsets $\{A_1, A_2, \dots\}$, it holds that

$$P\left[\bigcup_{n=1}^{\infty} A_n\right] = \sum_{n=1}^{\infty} P[A_n]$$

Probability Space

A probability space is a triplet Ω, F, P .

Probability Properties

$$P[A \cup B] = P[A] + P[B] - P[A \cap B] \quad (33)$$

$$P[A \cup B] \leq P[A] + P[B] \quad (34)$$

$$A \subseteq B \implies P[A] \leq P[B] \quad (35)$$

$$P[A|B] = \frac{P[A \cap B]}{P[B]} \quad (36)$$

Outcomes are statistically *independent* if $P(A|B) = P(A)$ (assuming $P(B) > 0$), or equivalently $P(A \cap B) = P(A)P(B)$.

Bayes Theorem states that for any two events A and B such that $P[A] > 0$ and $P[B] > 0$,

$$P[A|B] = \frac{P[B|A]P[A]}{P[B]} \quad (37)$$

The *Law of Total Probability* states that if $\{A_1, A_2, \dots, A_n\}$ is a partition of Ω , then for any $B \subseteq \Omega$,

$$P[B] = \sum_{i=1}^n P[B|A_i]P[A_i] \quad (38)$$

Random Variables

A *random variable* X is a function $X : \Omega \Rightarrow \Re$ that maps an outcome $\epsilon \in \Omega$ to a number $X(\epsilon)$ on the real line. We call it a variable because it has multiple states.

A *Bernoulli random variable* has a state of either 0 or 1. The probability of getting 1 is p and the probability of getting 0 is $1 - p$. We write

$$X \text{ Bernoulli}(p) \quad (39)$$

to say that X is drawn from a Bernoulli distribution with a parameter p . For a Bernoulli distribution,

$$E[X] = p \quad (40)$$

$$E[X^2] = p \quad (41)$$

$$\text{Var}[X] = p(1 - p) \quad (42)$$

A *Rademacher random variable* has two states, -1 and 1. The probability of getting each is 0.5.

A *binomial random variable* has a PMF of

$$p_X(k) = \binom{n}{k} p^k (1 - p)^{n-k}, k = 0, 1, \dots, n \quad (43)$$

where $0 < p < 1$ is the binomial parameter, and n is the total number of states. We write

$$X \text{ Binomial}(n, p) \quad (44)$$

to say that X is drawn from a binomial distribution with a parameter p of size n . If $X \text{ Binomial}(n, p)$, then

$$E[X] = np \quad (45)$$

$$E[X^2] = np(np + (1 - p)) \quad (46)$$

$$\text{Var}[X] = np(1 - p) \quad (47)$$

Let X be a *geometric random variable*. Then the PMF of X is

$$p_X(k) = (1 - p)^{k-1} p, k = 1, 2, \dots \quad (48)$$

We write

$$X \text{ Geometric}(p) \quad (49)$$

to say that X was drawn from a geometric distribution with a parameter p . If $X \text{ Geometric}(p)$ then

$$E[X] = \frac{1}{p} \quad (50)$$

$$E[X^2] = \frac{2}{p^2} - \frac{1}{p} \quad (51)$$

$$\text{Var}[X] = \frac{1 - p}{p^2} \quad (52)$$

Let X be a *Poisson random variable*. Then the PMF of X is

$$p_X(k) = \frac{\lambda^k}{k!} e^{-\lambda}, k = 0, 1, 2, \dots \quad (53)$$

where $\lambda > 0$ is the Poisson rate. We write $X \text{ Poisson}(\lambda)$ to say that X was drawn from a Poisson distribution with a parameter λ . If $X \text{ Poisson}(\lambda)$ then

$$E[X] = \lambda \quad (54)$$

$$E[X^2] = \lambda + \lambda^2 \quad (55)$$

$$\text{Var}[X] = \lambda \quad (56)$$

For small p and large n ,

$$\binom{n}{k} p^k (1-p)^{n-k} \approx \frac{\lambda^k}{k!} e^{-\lambda} \quad (57)$$

The *expectation* of a random variable X is

$$E[X] = \sum_{x \in X(\Omega)} x p_X(x) \quad (58)$$

The difference between $E[X]$ and the mean is that $E[X]$ is computed from the ideal histogram, while mean is computed from the empirical histogram. In general for any functions g and h ,

$$E[g(X)] = \sum_x g(x) p_X(x) \quad (59)$$

$$E[g(X) + h(X)] = E[g(X)] + E[h(X)] \quad (60)$$

$$E[cX] = cE[X] \quad (61)$$

$$E[X + c] = E[X] + c \quad (62)$$

The *variance* of a random variable X is

$$\text{Var}[X] = E[(X - \mu)^2] \quad (63)$$

or alternatively, the second moment minus the first moment squared.

$$E[X^2] - E[X]^2 \quad (64)$$

The *probability mass function* (PMF) $p_X(a)$ of a random variable X specifies the probability of obtaining a number $X(\epsilon) = a$. We denote a PMF as

$$p_X(a) = P[X = a] \quad (65)$$

PMFs are represented with histograms. A PMF should satisfy

$$\sum_{x \in X(\Omega)} p_X(x) = 1 \quad (66)$$

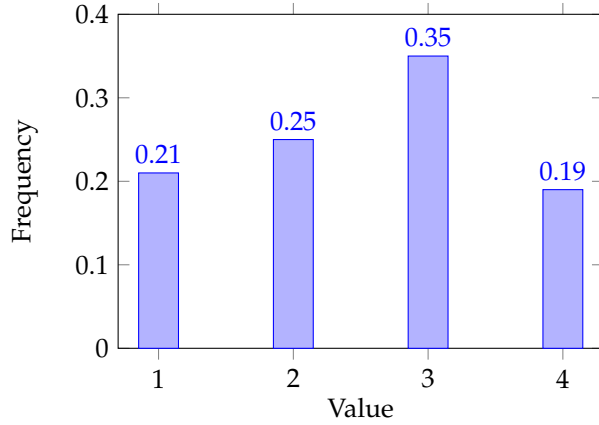


Figure 1: PMF

The *cumulative distribution function* is given by

$$F_X(x) = P[X \leq x] \quad (67)$$

$$= \sum_{u \leq x} p_X(u) \quad (68)$$

and represents the sum of every impulse of the PMF up to x .

Joint Distributions are higher-dimensional PDFs, PMFs, or CDFs. We write

$$f_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n) \equiv f_{\vec{X}}(\vec{x}) \quad (69)$$

The *joint PMF* of two random variables X and Y is notated by

$$p_{X,Y}(x, y) = P[X = x \text{ and } Y = y] \quad (70)$$

and represents the probability of both.

A *marginal PMF* is defined as

$$p_X(x) = \sum_{y \in \Omega_Y} p_{X,Y}(x, y) \quad (71)$$

or w.l.o.g.

$$p_Y(y) = \sum_{x \in \Omega_X} p_{X,Y}(x, y) \quad (72)$$

That is, it is the joint PMF summed over one of the variables.

If two random variables X and Y are independent, then

$$p_{X,Y} = p_X(x)p_Y(y) \quad (73)$$

$$f_{X,Y} = f_X(x)f_Y(y) \quad (74)$$

If a sequence of random variables X_1, X_2, \dots, X_N are independent, then their joint PDF (or joint PMF) can be factorized as

$$f_{X_1, X_2, \dots, X_N}(x_1, x_2, \dots, x_N) = \prod_{n=1}^N f_{X_n}(x_n) \quad (75)$$

The *joint CDF* of two random variables X and Y is the function $F_{X,Y}(x,y)$ such that

$$F_{X,Y}(x,y) = P[X \leq x \cap Y \leq y] \quad (76)$$

If X and Y are discrete, then

$$F_{X,Y}(x,y) = \sum_{y' \leq y} \sum_{x' \leq x} p_{X,Y}(x',y') \quad (77)$$

For two random variables X and Y , the *marginal CDF* is

$$F_X(x) = F_{X,Y}(x, \infty) \quad (78)$$

$$F_Y(y) = F_{X,Y}(\infty, y) \quad (79)$$

Let X and Y be two random variables. The *joint expectation* is

$$E[XY] = \sum_{y \in \Omega_Y} \sum_{x \in \Omega_X} xy \times p_{X,Y}(x,y) \quad (80)$$

If X and Y are discrete, then joint expectation is also called *correlation*.

This can be written in matrix form as

$$\begin{bmatrix} p_{X,Y}(x_1, y_1) & p_{X,Y}(x_1, y_2) & \dots & p_{X,Y}(x_1, y_N) \\ p_{X,Y}(x_2, y_1) & p_{X,Y}(x_2, y_2) & \dots & p_{X,Y}(x_2, y_N) \\ \vdots & \vdots & \ddots & \vdots \\ p_{X,Y}(x_N, y_1) & p_{X,Y}(x_N, y_2) & \dots & p_{X,Y}(x_N, y_N) \end{bmatrix} \quad (81)$$

then the joint expectation is

$$E[XY] = \sum_{i=1}^N \sum_{j=1}^N x_i y_j \times p_{X,Y}(x_i, y_j) \quad (82)$$

Let the matrix in Equation 81 be \mathbf{P} . Let

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} \quad (83)$$

$$\vec{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} \quad (84)$$

then

$$E[XY] = \begin{bmatrix} x_1 & x_2 & \dots & x_N \end{bmatrix} \begin{bmatrix} p_{X,Y}(x_1, y_1) & p_{X,Y}(x_1, y_2) & \dots & p_{X,Y}(x_1, y_N) \\ p_{X,Y}(x_2, y_1) & p_{X,Y}(x_2, y_2) & \dots & p_{X,Y}(x_2, y_N) \\ \vdots & \vdots & \ddots & \vdots \\ p_{X,Y}(x_N, y_1) & p_{X,Y}(x_N, y_2) & \dots & p_{X,Y}(x_N, y_N) \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} \quad (85)$$

$$= \vec{x}^T \mathbf{P} \vec{y} \quad (86)$$

$E[XY]$ is a weighted inner product between the states. \vec{x} and \vec{y} are the states of the random variables X and Y . Recalling that the magnitude of the inner product of \vec{a} and \vec{b} is $|\vec{a}||\vec{b}|\cos(\theta)$ and that cosine is bounded, we have

$$-1 \leq \frac{E[XY]}{\sqrt{E[X^2]}\sqrt{E[Y^2]}} \leq 1 \quad (87)$$

Notice that the correlation of X, Y is proportional to the covariance.

Reference

Series

$$\sum_{k=0}^n r^k = \frac{1 - r^{n+1}}{1 - r} \quad (1)$$

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6} \quad (2)$$

$$\sum_{k=1}^{\infty} kr^{k-1} = \frac{1}{(1-r)^2} \quad (3)$$

Combinatorics

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \quad (4)$$

$$(a+b)^n = \sum_{k=0}^n \binom{n}{k} a^{n-k} b^k \quad (5)$$

$$\binom{n}{k} + \binom{n}{k-1} = \binom{n+1}{k} \quad (6)$$

$$P(n, k) = \frac{n!}{(n-k)!} \quad (7)$$

where $P(n, k)$ is the number of ways to arrange k objects out of n (permutations).

$$C(n, k) = \binom{n}{k} = \frac{n!}{k!(n-k)!} \quad (8)$$

where $C(n, k)$ is the number of ways to choose k objects out of n (combinations).

Approximations

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \dots \quad (9)$$

$$= \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!} (x-a)^n \quad (10)$$

$$1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots = \sum_{k=0}^{\infty} \frac{x^k}{k!} \quad (11)$$

$$= e^x \quad (12)$$

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots \quad (13)$$

$$= \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} \quad (14)$$

$$\cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots \quad (15)$$

$$= \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} \quad (16)$$

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots \quad (17)$$

$$= \sum_{n=1}^{\infty} (-1)^{n+1} \frac{x^n}{n} \quad (18)$$

Calculus

$$\frac{d}{dx} \int_a^x f(t) dt = f(x) \quad (19)$$

$$\int_a^b f'(x) dx = f(b) - f(a) \quad (20)$$

$$\int f(g(x))g'(x) dx = \int f(u) du \quad (21)$$

$$\int u dv = uv - \int v du \quad (22)$$

$$\int \frac{1}{(x-a)(x-b)} dx = \frac{1}{b-a} \ln \left| \frac{x-a}{x-b} \right| + C \quad (23)$$

Linear Algebra

$$\vec{y} = \beta_1 \vec{x}_1 + \beta_2 \vec{x}_2 + \dots + \beta_N \vec{x}_N \quad (24)$$

$$\langle \vec{a}, \vec{b} \rangle = \vec{a} \vec{b}^T \quad (25)$$

$$= \sum_{i=1}^n a_i b_i \quad (26)$$

where $\langle \vec{a}, \vec{b} \rangle$ denotes the inner product of vectors \vec{a} and \vec{b} .

$$\|\vec{x}\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{1/p} \quad (27)$$

where $\|\vec{x}\|_p$ is the p -norm (or ℓ_p -norm) of vector \vec{x} .

$$\cos(\theta) = \frac{\langle \vec{a}, \vec{b} \rangle}{\|\vec{a}\|_2 \|\vec{b}\|_2} \quad (28)$$

where θ is the angle between vectors \vec{a} and \vec{b} .

$$\hat{\beta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \vec{y} \quad (29)$$

where $\hat{\beta}$ is the vector of least squares coefficients, \mathbf{X} is the data matrix, and \vec{y} is the target vector