DeMorgan's Laws:

- $\overline{A \cap B} = \overline{A} \cup \overline{B}$
- $\overline{A \cup B} = \overline{A} \cap \overline{B}$

Countable if elements can be put in one-to-one with $\mathbb{N} = \{1, 2, 3, \ldots\}$

Uncountable sets:

- $\mathbb{R} = (-\infty, \infty)$
- $[0, 1]$ and $(0, 1)$
- $(a, b), \forall a, b \in \mathbb{R}$ such that $a < b$

A family of sets $\{A_i, i \in I\}$ is a **partition of** $\mathcal{S}$ if it is disjoint and collectively exhaustive over $\mathcal{S}$.

Axioms of Probability

1. $\mathcal{P}(A) \geq 0, \forall A \in \mathcal{F}(\mathcal{S}))$
2. $\mathcal{P}(\mathcal{S}) = 1$
3. $\mathcal{P}(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathcal{P}(A_i)$

Probability Equations

1. $\mathcal{P}(A \cup B) = \mathcal{P}(A) + \mathcal{P}(B) - \mathcal{P}(A \cap B)$
2. $\mathcal{P}(A|B) = \frac{\mathcal{P}(A \cap B)}{\mathcal{P}(B)} = \frac{\mathcal{P}(B|A)\mathcal{P}(A)}{\mathcal{P}(B)}$

Given $\mathcal{S}$ and a family of subsets $\mathcal{G} = \{A_i, i \in I\}$ of $\mathcal{S}$, the **σ-field generated by** $\mathcal{G}$, denoted $\sigma(\mathcal{G})$, is the smallest σ field containing all the subsets in $\mathcal{G}$.
By "smallest" σ-field, we mean that for any σ-field $\mathcal{F}_0$ containing all the sets in $\mathcal{G}$:

$$\sigma(\mathcal{G}) \subset \mathcal{F}_0$$

Given $\mathbb{R}$, the **Borel field of** $\mathbb{R}$ is defined as the σ-field generated by the family of all open intervals

$$\mathcal{G} = \{(a, b) : \forall (a, b) \in \mathbb{R} \text{ such that } a < b\}.$$

We denote the Borel field of $\mathbb{R}$ by $\mathcal{B}(\mathbb{R})$.

---

pmfs
uniform
$p(\omega) = \frac{1}{n}, \forall \omega \in \mathcal{S}$
binomial
$p(k) = \binom{n}{k} a^k (1-a)^{n-k}$
geometric
$p(k) = (1-a)a^k, a \in (0, 1)$
poisson
$p(k) = \frac{\lambda^k e^{-k}}{k!}; k = 0, 1, 2, \ldots; \lambda > 0$

Properties of the pdf:

1. $f(r) \geq 0, \forall r \in \mathbb{R}$
2. $\int_{-\infty}^{\infty} f(r)dr = 1$

Given a valid pdf, we can get a valid probability measure $\mathcal{P}(\cdot)$:

$$\mathcal{P}(A) = \int_A f(r)dr = \int_{-\infty}^{\infty} \cdot 1_A(r)dr$$

where

$$1_A(r) = \begin{cases} 1, r \in A \\ 0, r \notin A \end{cases}$$

is called the **indicator function** of the set A.

---

pdfs
uniform
$f(r) = \frac{1}{b-a} 1_{[a,b]}(r)$
exponential
$f(r) = \lambda e^{-\lambda r} \cdot 1_{[0,\infty]}(r), \lambda > 0$
$\hookrightarrow = \begin{cases} \lambda e^{-\lambda r}, r \geq 0 \\ 0, r < 0 \end{cases}$
gaussian
$f(r) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{-(r-\mu)^2}{2\sigma^2}\right), r \in \mathbb{R}, \mu \in \mathbb{R}, \sigma > 0$

Conditional Probability

$$\mathcal{P}(A|B) = \frac{\mathcal{P}(A \cap B)}{\mathcal{P}(B)}$$

Bayes Formula

$$\mathcal{P}(A|B) = \frac{\mathcal{P}(B|A)\mathcal{P}(A)}{\mathcal{P}(B)}$$

Total Probability Law

$$\mathcal{P}(B) = \sum_{i=1}^{n} \mathcal{P}(B|A_i)\mathcal{P}(A_i)$$

Bayes Theorem

$$\mathcal{P}(A_m|B) = \frac{\mathcal{P}(B|A_m)\mathcal{P}(A_m)}{\sum_{i=1}^{n} \mathcal{P}(B|A_i)\mathcal{P}(A_i)}$$

Events A and B are **statistically independent** if and only if

$$\mathcal{P}(A \cap B) = \mathcal{P}(A)\mathcal{P}(B)$$

Bernoulli Trials

$$p_n(k) = \mathcal{P}(\mathcal{B}_k) = \binom{n}{k} p^k (1-p)^{n-k}$$

where $p = \mathcal{P}_0(A)$.

Properties of the cdf

$$F_{\mathbb{X}}(\alpha) = \mathcal{P}_x((-\infty, \alpha)) = \mathcal{P}(\{\mathbb{X} \leq \alpha\})$$

where

$$\{\mathbb{X} \leq \alpha\} = \{\omega \in \mathcal{S} : \mathbb{X}(\omega) \leq \alpha\} \in \mathcal{F}$$

Properties:

1. $F_{\mathbb{X}}(\infty) = 1$ and $F_{\mathbb{X}}(-\infty) = 0$
2. If $x_1 < x_2$, then $F_{\mathbb{X}}(x_1) \leq F_{\mathbb{X}}(x_2)$
3. $\mathcal{P}(\{\mathbb{X} > \alpha\}) = 1 - F_{\mathbb{X}}(\alpha)$
4. If $x_1 < x_2$, then $\mathcal{P}(\{x_1 < \mathbb{X} \leq x_2\}) = F_{\mathbb{X}}(x_2) - F_{\mathbb{X}}(x_1)$
5. $\mathcal{P}(\{\mathbb{X} = x_0\}) = F_{\mathbb{X}}(x_o) - F_{\mathbb{X}}(x_0^-)$ where $F_{\mathbb{X}}(x_0^-) = \lim_{\epsilon \downarrow 0} F_{\mathbb{X}}(x_0 - \epsilon)$

Properties of the pdf

$$f_{\mathbb{X}}(x) = \frac{dF_{\mathbb{X}}(x)}{dx}$$

Properties:

1. $f_{\mathbb{X}}(x) \geq 0, \forall x \in \mathbb{R}$
2. $F_{\mathbb{X}}(x) = \int_{-\infty}^{x} f_{\mathbb{X}}(\alpha)d\alpha$
3. $\int_{\infty}^{\infty} f_{\mathbb{X}}(x)dx = F_{\mathbb{X}}(\infty) - F_{\mathbb{X}}(-\infty) = 1$
4. $\mathcal{P}(\{x_1 < \mathbb{X} \leq x_2\}) = \int_{x_1}^{x_2} f_{\mathbb{X}}(x)dx = F_{\mathbb{X}}(x_2) - F_{\mathbb{X}}(x_1)$

Dirac δ-functions: $\delta(x)$

- $\delta(x) = 0, \forall x \neq 0$
- $\int_{\infty}^{\infty} \delta(x)dx = \int_{-\epsilon}^{\epsilon} \delta(x)dx = 1, \forall \epsilon > 0$

Sifting Property of Dirac δ-functions

$$\int_{\infty}^{\infty} g(x)\delta(x - x_0)dx = g(x_0)$$

Random Variable Forms
gaussian

$$f_{\mathbb{X}}(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp[\frac{-(x-\mu)^2}{2\sigma^2}], \forall x \in \mathbb{R}$$

n.b.

$$F_{\mathbb{X}}(x) = \int_{-\infty}^{x} f_{\mathbb{X}}(\alpha)d\alpha = \Phi(\frac{x-\mu}{\sigma})$$

$$\text{where } \Phi(r) = \int_{-\infty}^{r} \frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}} dz$$

So if $\mathbb{X}$ is a Gaussian RV with parameters $\mu \in \mathbb{R}$ and $\sigma > 0$, then

$$\mathcal{P}(\{a < \mathbb{X} \leq b\}) = \Phi(\frac{b-\mu}{\sigma}) - \Phi(\frac{a-\mu}{\sigma})$$

uniform

$$\mathbb{X} = \mathrm{u}[a, b], a < b$$

$$f_{\mathbb{X}}(x) = \frac{1}{b-a} \cdot 1_{[a,b]}(x)$$

$$F_{\mathbb{X}}(x) = \int_{\infty}^{x} f_{\mathbb{X}}(\alpha)d\alpha$$

exponential
An RV $\mathbb{X}$ with pdf of the form

$$f_{\mathbb{X}}(x) = \alpha e^{-\alpha x} \cdot 1_{[0,\infty]}(x)$$

where $\alpha > 0$ is called an exponential RV with parameter α.

$$F_{\mathbb{X}}(\alpha) = \int_{-\infty}^{\alpha} f_{\mathbb{X}}(x)dx$$

binomial
Discrete RV taking on values in the set $\{0, 1, 2, \cdots, n\} \subset \mathbb{R}$ with pmf

$$\mathcal{P}_{\mathbb{X}}(k) = \binom{n}{k} p^k (1-p)^{n-k},$$

$$F_{\mathbb{X}}(x) = \sum_{k=0}^{m(x)} \binom{n}{k} p^k (1-p)^{n-k}$$

where $m(x) \leq x < m(x) + 1$,

$$f_{\mathbb{X}}(x) = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} \delta(x - k)$$

Conditional cdf and pdf

$$F_{\mathbb{X}}(x|M) = \mathcal{P}(\{\mathbb{X} \leq x\}|M) = \frac{\mathcal{P}(\{\mathbb{X} \leq x\} \cap M)}{\mathcal{P}(M)}$$

$$f_{\mathbb{X}}(x|M) = \frac{dF_{\mathbb{X}}(x|M)}{dx}$$

Total Probability Law

$$F_{\mathbb{X}}(x) = F_{\mathbb{X}}(x|A_1)P(A_1) + \cdots + F_{\mathbb{X}}(x|A_n)P(A_n)$$

$$f_{\mathbb{X}}(x) = f_{\mathbb{X}}(x|A_1)P(A_1) + \cdots + f_{\mathbb{X}}(x|A_n)P(A_n)$$

Bayes Formula

$$\mathcal{P}(A|\{\mathbb{X} \leq x\}) = \frac{F_{\mathbb{X}}(x|A)\mathcal{P}(A)}{F_{\mathbb{X}}(x)}$$

Now consider $\mathcal{P}(A|B) = \frac{\mathcal{P}(B|A)\mathcal{P}(A)}{\mathcal{P}(B)}$ when $B = \{x_1 < \mathbb{X} \leq x_2\}$:

$$\mathcal{P}(A|\{x_1 < \mathbb{X} \leq x_2\}) =$$
$$\hookrightarrow = \frac{F_\mathbb{X}(x_2|A) - F_\mathbb{X}(x_1|A)\mathcal{P}(A)}{F_\mathbb{X}(x_2) - F_\mathbb{X}(x_1)}$$

For $\mathbb{X} = x$,

$$\mathcal{P}(A|\{\mathbb{X} = x\}) = \frac{f_\mathbb{X}(x|A)}{f_\mathbb{X}(x)}\mathcal{P}(A)$$
$$\mathcal{P}(A) = \int_{-\infty}^{\infty} \mathcal{P}(A|\{\mathbb{X} = x\})f_\mathbb{X} dx$$

Bayes' Theorem

$$f_\mathbb{X}(x|A) = \frac{\mathcal{P}(A|\{\mathbb{X} = x\})f_\mathbb{X}(x)}{\int_{-\infty}^{\infty} \mathcal{P}(A|\{\mathbb{X} = \alpha\})f_\mathbb{X}(\alpha)d\alpha}$$

Functions of RVs
For $\mathbb{Y} = g(\mathbb{X})$ to be measurable, $g(\cdot)$ must satisfy:

1. The domain of $g(\cdot)$ must contain the range space of $\mathbb{X}$.
2. For each $y \in \mathbb{R}$, the set $R_y = \{x \in \mathbb{R} : g(x) \leq y\}$ must be a Borel set.
3. The events $\{g(\mathbb{X}) = \pm\infty\}$ must have probability 0.

Direct pdf Method
Suppose $\mathbb{Y} = g(\mathbb{X})$, where $G : \mathbb{R} \to \mathbb{R}$ such that the inverse $g^{-1}(\cdot)$ exists, and assume that

$$\frac{dx}{dy} = \frac{dg^{-1}(y)}{dy}$$

exists. Then

$$f_\mathbb{Y}(y) = f_\mathbb{X}(x(y)) \cdot \left|\frac{dx(y)}{dy}\right|$$

where $x(y) = g^{-1}(y)$.

Integration by parts

$$\int uv'du = uv - \int vu'du$$

The **mean** or **expected value** of a RV $\mathbb{X}$ with pdf $f_\mathbb{X}(x)$ is

$$E[\mathbb{X}] = \int_{-\infty}^{\infty} x f_\mathbb{X}(x)dx$$

The definition above applies to discrete RVs if we write their pdf using δ-functions:

$$f_\mathbb{X}(x) = \sum_k p_\mathbb{X}(x_k)\delta(x - x_k) = \sum_k p_k \delta(x - x_k)$$
$$E[\mathbb{X}] = \sum_k x_k p_\mathbb{X}(x_k) = \sum_k k \cdot p_\mathbb{X}(k)$$

If $\mathbb{Y} = g(\mathbb{X})$ is a RV, then

$$E[\mathbb{Y}] = E[g(\mathbb{X})] = \int_{\infty}^{\infty} g(x)f_\mathbb{X}(x)dx.$$

Conditional mean of an RV

$$E[\mathbb{X}|M] = \int_{-\infty}^{\infty} x f_\mathbb{X}(x|M)dx \text{ (continuous)}$$
$$E[\mathbb{X}|M] = \sum_k x_k p_\mathbb{X}(x_k|M) \text{ (discrete)}$$

Expected value of function $g(\mathbb{X})$

$$E[g(\mathbb{X})] = \int_{-\infty}^{\infty} g(x)f_\mathbb{X}(x)dx = \sum_k g(x_k)p_\mathbb{X}(x_k)$$

Linearity of Expectation
Let $g_1(\mathbb{X})$ and $g_2(\mathbb{X})$ be two function of a RV $\mathbb{X}$ and let α and β be two constant ($\alpha, \beta \in \mathbb{R}$). Then

$$E[\alpha g_1(\mathbb{X}) + \beta g_2(\mathbb{X})] = \alpha E[g_1(\mathbb{X})] + \beta E[g_2(\mathbb{X})]$$

Variance of RV $\mathbb{X}$

$$\text{var}(\mathbb{X}) = E[(\mathbb{X} - \overline{\mathbb{X}})^2] = \int_{-\infty}^{\infty} (x - \overline{\mathbb{X}})^2 f_\mathbb{X}(x)dx$$

where $\overline{\mathbb{X}} = E[\mathbb{X}]$.
The positive square root of the variance of $\mathbb{X}$ is called the **standard deviation** of $\mathbb{X}$:

$$\text{StDev}(\mathbb{X}) = \sigma_x = \sqrt{\text{var}(\mathbb{X})}$$

n.b.:

$$\text{var}(\mathbb{X}) = E[(\mathbb{X} - \overline{\mathbb{X}})^2] = E[\mathbb{X}^2] - (E[\mathbb{X}])^2$$

Means and variances of RV types
gaussian
$E[\mathbb{X}] = \mu$
$\text{var}(\mathbb{X}) = \sigma^2$
poisson
$E[\mathbb{X}] = \mu$
$\text{var}(\mathbb{X}) = \mu$

The Characteristic Function

$$\Phi_\mathbb{X}(\omega) = E[e^{i\omega\mathbb{X}}], \omega \in \mathbb{R}$$

If $f_\mathbb{X}(x)$ is the pdf of $\mathbb{X}$, then

$$\Phi_\mathbb{X}(\omega) = \int_{-\infty}^{\infty} e^{i\omega x} f_\mathbb{X}(x)dx$$
$$\Phi_\mathbb{X}(\omega) : \mathbb{R} \to \mathbb{C}$$

Euler's formula:

$$e^{i\omega x} = \cos(\omega x) + i\sin(\omega x)$$
$$e^{i\pi} + 1 = 0$$

If $\mathbb{X}$ is a RV with $\Phi_\mathbb{X}(\omega)$, and $\mathbb{Y} = a\mathbb{X} + b$ then

$$\Phi_\mathbb{Y}(\omega) = e^{i\omega b}\Phi_\mathbb{X}(a\omega)$$

Moment generating function

$$\phi_\mathbb{X}(s) = E[e^{s\mathbb{X}}] = \int_{-\infty}^{\infty} f_\mathbb{X}(x)e^{sx}dx$$

Moment Theorem
Given a RV $\mathbb{X}$ with mgf $\phi_\mathbb{X}(s)$, the n-th moment of $\mathbb{X}$ is given by

$$E[\mathbb{X}^n] = \frac{d^n\phi_\mathbb{X}(s)}{ds^n}|_{s=0} = \phi_\mathbb{X}^{(n)}(0)$$

The **joint cdf** of two RVs is the probability of the event $\{\mathbb{X} \leq x\} \cap \{\mathbb{Y} \leq y\}$:

$$F_{\mathbb{XY}}(x, y) = \mathcal{P}(\{\mathbb{X} \leq x\} \cap \{\mathbb{Y} \leq y\})$$

Properties:

1. $F_{\mathbb{XY}}(-\infty, y) = 0$ and $F_{\mathbb{XY}}(x, -\infty) = 0$, $F_{\mathbb{XY}}(\infty, y) = F_\mathbb{Y}(y)$ and $F_{\mathbb{XY}}(x, \infty) = F_\mathbb{X}(x)$, $F_{\mathbb{XY}}(\infty, \infty) = 1$
2. $\mathcal{P}(\{x_1 < \mathbb{X} \leq x_2\} \cap \{\mathbb{Y} \leq y\}) = F_{\mathbb{XY}}(x_2, y) - F_{\mathbb{XY}}(x_1, y)$
3. $\mathcal{P}(\{x_1 < \mathbb{X} \leq x_2\} \cap \{y_1 < \mathbb{Y} \leq y_2\}) = F_{\mathbb{XY}}(x_2, y_2) - F_{\mathbb{XY}}(x_2, y_1)$

Joint pdf

$$f_{\mathbb{XY}}(x, y) = \frac{\partial^2 F_{\mathbb{XY}}(x, y)}{\partial x \partial y}$$

Properties:

1. $f_{\mathbb{XY}}(x, y) \geq 0$
2. $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{\mathbb{XY}}(x, y)dxdy = 1$
3. $\int_{-\infty}^{y}\int_{-\infty}^{x} f_{\alpha\beta}(x, y)d\alpha d\beta = F_{\mathbb{XY}}(x, y)$
4. $\mathcal{P}(\{(\mathbb{X}, \mathbb{Y}) \in D\}) = \int\int_{\mathbb{R}^2} f_{\mathbb{XY}}(x, y) \cdot 1_D((x, y))dxdy$

Marginal pdfs

$$f_\mathbb{X}(x) = \int_{-\infty}^{\infty} f_{\mathbb{XY}}(x, y)dy$$
$$f_\mathbb{Y}(y) = \int_{-\infty}^{\infty} f_{\mathbb{XY}}(x, y)dx$$

Jointly Gaussian RVs

$$f_{\mathbb{XY}}(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1 - r^2}} \cdot$$
$$\exp\left[\cdots \frac{(x - \mu_x)(y - \mu_y)}{\sigma_x\sigma_y} \cdots\right]$$

where $\mu_x, \mu_y \in \mathbb{R}$, $\sigma_x\sigma_y > 0$, and $-1 \leq r \leq 1$.

Statistically Independent RVs
Two RVs are statistically independent if the events $\{\mathbb{X} \in A\}$ and $\{\mathbb{Y} \in B\}$ are statistically independent for all $A, B \in \mathcal{B}(\mathbb{R})$.

$$F_{\mathbb{XY}}(x, y) = \mathcal{P}(\{\mathbb{X} \in A\} \cap \{\mathbb{Y} \in B\}) = F_\mathbb{X}(x) \cdot F_\mathbb{Y}($$
$$f_{\mathbb{XY}}(x, y) = f_\mathbb{X}(x) \cdot f_\mathbb{Y}(y)$$

If $\mathbb{X}$ and $\mathbb{Y}$ are two j-dist, independent RVs then the pdf of their sum $\mathbb{Z} = \mathbb{X} + \mathbb{Y}$ is given by the convolution:

$$f_\mathbb{Z}(z) = (f_\mathbb{X} * f_\mathbb{Y})(z)$$
$$f_\mathbb{Z}(z) = \int_{-\infty}^{\infty} f_\mathbb{Y}(y)f_\mathbb{X}(z - y)dy$$

Direct joint density determination

$$f_{\mathbb{ZW}}(z, w) = f_{\mathbb{XY}}(x(z, w), y(z, w)) \cdot \left|\frac{\partial(x, y)}{\partial(z, w)}\right|$$

where the Jacobian is the determinant

$$\frac{\partial(x, y)}{\partial(z, w)} = \frac{\partial x}{\partial z} \cdot \frac{\partial y}{\partial w} - \frac{\partial y}{\partial z} \cdot \frac{\partial x}{\partial w}.$$

Joint Moments

$$E[\mathbb{Z}] = E[g(\mathbb{X}, \mathbb{Y})]$$
$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x, y)f_{\mathbb{XY}}(x, y)dxdy$$

Correlation

$$\text{corr}(\mathbb{X}, \mathbb{Y}) = E[\mathbb{X}\mathbb{Y}]$$

Covariance

$$\text{cov}(\mathbb{X}, \mathbb{Y}) = E[(\mathbb{X} - \overline{\mathbb{X}})(\mathbb{Y} - \overline{\mathbb{Y}})]$$

Correlation coefficient

$$r_{\mathbb{X}\mathbb{Y}} = \frac{\text{cov}(\mathbb{X}, \mathbb{Y})}{\sigma_x \sigma_y} = \frac{E[\mathbb{X}\mathbb{Y}] - E[\mathbb{X}] \cdot E[\mathbb{Y}]}{\sigma_x \sigma_y}$$

Two RVs are **uncorrelated** if their covariance is equal to zero. This is true if any of the following are true:

1. $\text{cov}(\mathbb{X}, \mathbb{Y}) = 0$
2. $r_{\mathbb{X}\mathbb{Y}} = 0$
3. $E[\mathbb{X}\mathbb{Y}] = E[\mathbb{X}] \cdot E[\mathbb{Y}]$

Two RVs are **orthogonal** if $E[\mathbb{X}\mathbb{Y}] = 0$.

Joint characteristic functions

$$\Phi_{\mathbb{X}\mathbb{Y}}(\omega_1, \omega_2) = E[e^{i(\omega_1 \mathbb{X} + \omega_2 \mathbb{Y})}]$$
$$= \int\int e^{+i(\omega_1 x + \omega_2 y)} f_{\mathbb{X}\mathbb{Y}}(x, y) dx dy$$

Convolution Theorem

Let $\mathbb{X}$ and $\mathbb{Y}$ be two j-dist, statistically independent RVs and let $\mathbb{Z} = \mathbb{X} + \mathbb{Y}$, then

$$\Phi_{\mathbb{Z}}(\omega) = \Phi_{\mathbb{X}}(\omega) \cdot \Phi_{\mathbb{Y}}(\omega) = E[e^{i\omega\mathbb{X}}] \cdot E[e^{i\omega\mathbb{Y}}]$$

Joint moment generating function

$$\phi_{\mathbb{X}\mathbb{Y}}(s_1, s_2) = E[e^{s_1 \mathbb{X} + s_2 \mathbb{Y}}]$$

Moment Theorem

$$E[\mathbb{X}^j \cdot \mathbb{Y}^k] = \phi_{\mathbb{X}\mathbb{Y}}^{(j,k)}(0, 0)$$

Conditional Distributions

$$F_{\mathbb{X}\mathbb{Y}}(x, y|M) = \mathcal{P}(\{\mathbb{X} \le x\} \cap \{\mathbb{Y} \le y\}|M)$$
$$= \frac{\mathcal{P}(\{\mathbb{X} \le x\} \cap \{\mathbb{Y} \le y\} \cap M)}{\mathcal{P}(M)}$$

$$f_{\mathbb{Y}}(y|\{\mathbb{X} = x\}) = \frac{f_{\mathbb{X}\mathbb{Y}}(x, y)}{f_{\mathbb{X}}(x)}$$

Random Vectors

The vector of j-dist RVs

$$\underline{\mathbb{X}} = (\mathbb{X}_1, \cdots, \mathbb{X}_n)$$

is a random vector.

$$F_{\underline{\mathbb{X}}}(\underline{x}) = F_{\mathbb{X}_1 \cdots \mathbb{X}_n}(x_1, \cdots, x_n)$$
$$f_{\underline{\mathbb{X}}}(\underline{x}) = f_{\mathbb{X}_1 \cdots \mathbb{X}_n}(x_1, \cdots, x_n)$$

Central Limit Theorem

Let $\{\mathbb{X}_n\}$ be a sequence of iid RVs with mean μ and variance $\sigma^2 < \infty$. Define

$$\mathbb{Z}_n = \frac{(\mathbb{X}_1 + \mathbb{X}_2 + \cdots + \mathbb{X}_n) - n\mu}{\sigma\sqrt{n}}$$

Then $\{Z_n\}$ converges in distribution to a RV $\mathbb{Z}$ that is Gaussian with mean 0 and variance 1. i.e.,

$$F_{\mathbb{Z}_n}(z) \to \Phi(z) \text{ as } n \to \infty, \forall z \in \mathbb{R}$$

Stochastic Processes

Instead of mapping each outcome $\omega \in \mathcal{S}$ to a number $\mathbb{X}(\omega)$, we map it to a function of time $\mathbb{X}(t, \omega)$.
A **random process** is a family $\{\mathbb{X}(t); t \in \mathbb{R}\}$ of RVs defined on $(\mathcal{S}, \mathcal{F}, \mathcal{P})$ and indexed by t.
For a fixed outcome $\omega_0 \in \mathcal{S}$, the time function $\mathbb{X}(\cdot) = \mathbb{X}(\cdot, \omega_0)$ is called the **sample path** of the random process $\mathbb{X}(t)$ corresponding to ω_0.

---

n-th order stationary

A random process $\mathbb{X}(t)$ is called n-th order stationary if all n-th order cdfs or pdfs are invariant to time shifts in the origin.

Strict-sense stationary (SSS)

A random process $\mathbb{X}(t)$ is called stationary, or SSS, if $\mathbb{X}(t)$ is n-th order stationary for all $n = 1, 2, 3, \cdots$

Wide-sense stationary (WSS)

A random process $\mathbb{X}(t)$ is called **WSS** if:

1. $E[\mathbb{X}(t)] = \overline{\mathbb{X}} = \text{constant}$
2. $E[\mathbb{X}(t_1) \cdot \mathbb{X}(t_2)] = R(t_1 - t_2)$

Note:

1. If a random process $\mathbb{X}(t)$ is first and second order stationary, and if $E[\mathbb{X}(t_1)]$ and $E[\mathbb{X}(t_1)\mathbb{X}(t_2)]$ exits, then $\mathbb{X}(t)$ is WSS. The converse is **not** true.
2. If a RP $\mathbb{X}(t)$ is SSS, it is also WSS if $E[\mathbb{X}(t_1)]$ and $E[\mathbb{X}(t_1)\mathbb{X}(t_2)]$ exist.
3. If $\mathbb{X}(t)$ is WSS, $E[\mathbb{X}(t_1)\mathbb{X}(t_2)]$, $\text{cov}(\mathbb{X}(t_1), \mathbb{X}(t_2))$ will depend only on the time difference $t_1 - t_2$.

Mean of RP $\mathbb{X}(t)$

$$\mu_{\mathbb{X}}(t) = E[\mathbb{X}(t)]$$

Autocorrelation function

$$R_{xx}(t_1, t_2) = E[\mathbb{X}(t_1)\mathbb{X}(t_2)]$$

Autocovariance function

$$C_{xx}(t_1, t_2) = E[(\mathbb{X}(t_1) - \mu_x(t_1))(\mathbb{X}(t_2) - \mu_x(t_2))]$$

Power spectral density

If $\mathbb{X}(t)$ is WSS with autocorrelation function $R_x(\tau)$, then the **power spectral density** of $\mathbb{X}(t)$ is defined as

$$S_{xx}(\omega) = \int_{-\infty}^{\infty} R_x(\tau) e^{-i\omega\tau} d\tau$$

$S_{xx}(\omega)$ is a measure of the average distribution of signal power in frequency for the RP $\mathbb{X}(t)$. This is the Fourier transform of the autocorrelation function for a WSS process.

A linear system $L[\cdot]$ is a transformation rule satisfying the following two properties:

1. $L[\mathbb{X}_1(t) + \mathbb{X}_2(t)] = L[\mathbb{X}_1(t)] + L[\mathbb{X}_2(t)]$
2. $L[\mathbb{A} \cdot \mathbb{X}(t)] = \mathbb{A} \cdot L[\mathbb{X}(t)]$

A (linear) system is **time-invariant** if, given response $y(t)$ to input $x(t)$, it has response $y(t + c)$ for input $x(t + c)$ for all $c \in \mathbb{R}$. If we put a random process $\mathbb{X}(t)$ into an LTI system, we get a random process $\mathbb{Y}(t)$ out:

$$\mathbb{Y}(t) = \mathbb{X}(t) * h(t) = \int_{-\infty}^{\infty} \mathbb{X}(t - \alpha) h(\alpha) d\alpha$$

If $\mathbb{X}(t)$ is WSS and its the input to a stable LTI system with impulse response $h(t)$, the the power spectral density of the output $\mathbb{Y}(t)$ is

$$S_{yy}(\omega) = Sxx(\omega)|H(\omega)|^2$$

where

$$H(\omega) = \int_{-\infty}^{\infty} h(t) e^{-i\omega t} dt$$

White noise process

$$C_{WW}(t_1, t_2) = 0, \forall t_1 \neq t_2$$

All WSS white noise processes have $R_{WW}(t_1, t_2) = r_0 \cdot \delta(t_1 - t_2)$ where $r_0 = \text{constant} > 0$

Exam topics

- LTI system, impulse response, compute Fourier transform