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 $\mathbb{N} = \{1, 2, 3, \dots\}$

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 $\{\mathcal{A}_i, i \in I\}$ 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 $\mathcal{G} = \{\mathcal{A}_i, i \in I\}$ 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
$p(k) = \frac{\lambda^k e^{-k}}{k!}; k = 0, 1, 2, \dots; \lambda > 0$

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.
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 $B = \{x_1 < \mathbb{X} \le x_2\}$:

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

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 $R_y = \{x \in \mathbb{R} : g(x) \le y\}$ 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:

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

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 $\mathbb{X}$

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

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

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

n.b.:

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

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