Sample Space ($\Omega$): Set of all possible outcomes
Events: a set of outcomes / a subset of $\Omega$

Countably infinite: if exists 1-1 correspondence $f: \Omega \to \mathbb{N}$

Disjoint: $S \cap T = \phi$
Mutually exclusive: $S_i \cap S_j = \phi$, $\forall i, j$, $i \neq j$

Complement: $S^c = \{x \in \Omega : x \notin S\}$
Subset: $S \subseteq T \iff$ For every $x \in S$, $x \in T$
Equal: $S = T \iff S \subseteq T$ and $T \subseteq S$

$\bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} S_n = \{x \in \Omega : x$ appears in infinitely many $S_n\}$
$\bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} S_n = \{x \in \Omega : x$ appears in all $S_n$ after some $k\}$

$\cup$ increasing $\downarrow$ $A_1 \cap A_2 \cap A_3 \dots \subseteq A_1 \subseteq A_1 \cup A_2 \cup A_3 \dots$ $\downarrow$ decreasing $\cap$
$A_2 \cap A_3 \dots \subseteq A_2 \subseteq A_2 \cup A_3 \dots$

$\bigcup_{n=1}^{\infty} S_n = \sup S_n$ $\Rightarrow$ $\bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} S_n = \lim_{n\to\infty} \sup S_n$
$\bigcap_{n=1}^{\infty} S_n = \inf S_n$ $\qquad$ $\bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} S_n = \lim_{n\to\infty} \inf S_n$

pf. $\underset{(S)}{\bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} S_n} = \underset{(T)}{\{x \mid x \in S_k . \forall k \text{ except finitely many } k\}}$
($S \subseteq T$): Let $x$ be an element of $S$. i.e. $x \in S$. Then there must exist some $k$ such that $x \in \bigcap_{n=k}^{\infty} S_n$, then $x \in S_k, S_{k+1}, S_{k+2}, \dots \Rightarrow x \in T$
($T \subseteq S$): pick some $y \in T$. so there must exist some $m \in \mathbb{N}$ s.t $y \in S_i$ $\forall i \geq m \Rightarrow y \in \bigcap_{n=i}^{\infty} S_n$ $\forall i \geq m \Rightarrow y \in \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} S_n$

pf. $\underset{(S)}{\bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} S_n} = \underset{(T)}{\{x \mid x \in S_k \text{ for infinitly many } k\}}$
($S \subseteq T$): Let $x$ be an element of $S$. i.e. $x \in S$. then $x \in \bigcup_{n=k}^{\infty} S_n, \forall k$. This implies that $x \in T$:
(assume $x \notin T \Rightarrow \exists m$ s.t $x \in S_{i_1}, S_{i_2}, \dots, S_{i_m}$ $i_1 < i_2 < \dots < i_m$, so $x \notin \bigcup_{n=i_m+1}^{\infty} S_n$. $\rightarrow\!\leftarrow$)
($T \subseteq S$): Let $y$ be an element of $T$, i.e. $y \in T$, then there must exist a countably infinite seq. $\{i_m\}$ s.t $y \in S_m$ $\forall m \in \mathbb{N}$. This implies that $y \in \bigcup_{n=k}^{\infty} S_n . \forall k \in \mathbb{N}. \Rightarrow y \in \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} S_n$

De Morgan's Laws
$(\bigcup_n S_n)^c = \bigcap_n S_n^c$
$(\bigcap_n S_n)^c = \bigcup_n S_n^c$

## 3 Axioms of Probability

A probability assignment is valid if

1. $P(A) \geq 0$ for any event $A$ $\quad$ 2. $P(\Omega) = 1$
3. $A_1, A_2 \dots$ are mutually exclusive, then $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$

$P(\phi) = 0$ $\quad (A_1 = \Omega, A_2 = A_3 \dots = \phi)$

$A_1, \dots, A_n$ are disjoint events, then $P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i)$

$P(A) \leq 1$ for any $A$ ($A \cup A^c = \Omega$, $A$ and $A^c$ are disjoint)

$P(A^c) = 1 - P(A)$ $\quad (A_1 = A, A_2 = A^c, A_3 = A_4 \dots = \phi)$

$P(A) = P(A-B) + P(A \cap B)$ $\quad (A_1 = A-B, A_2 = A \cap B, A_3 = A_4 \dots = \phi)$

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$ $\quad \left(\begin{array}{l} P(A \cup B) = P(A) + P(B-A) \\ P(B) = P(B-A) + P(A \cap B) \end{array}\right)$

## Union Bound

$P(\bigcup_{n=1}^{N} A_n) \leq \sum_{n=1}^{N} P(A_n)$

for any events $A_1, \dots, A_N$

(pf) step 1. $N=2$ $P(A_1 \cup A_2) = P(A_1 \cup (A_2 - A_1))$
$= p(A_1) + p(A_2 - A_1) \leq p(A) + p(A_2)$

step 2. assume $N=k$ is true. then when $N=k+1$.
$P(\bigcup_n^{k+1} A_n) = P(\bigcup_n^{k} A_n \cup A_{k+1}) \leq P(\bigcup_n^{k} A_n) + p(A_{k+1})$ step 1.
assumption $\leq \sum_{n=1}^{k} P(A_n) + p(A_{k+1})$

## Continuity of Probability Function

Probability function $P(\cdot)$ is a function of events (needs to satisfy the 3 axioms)

$E_1, E_2 \dots$ is increasing if $E_1 \subseteq E_2 \subseteq \dots \subseteq E_n \subseteq E_{n+1} \subseteq \dots$
decreasing if $E_1 \supseteq E_2 \supseteq \dots \supseteq E_n \supseteq E_{n+1} \supseteq \dots$

Theorem For any increasing / decreasing sequence of events $E_1, E_2 \dots$, we have
$$\lim_{n\to\infty} P(E_n) = P(\lim_{n\to\infty} E_n)$$

(pf) $P(\lim_{n\to\infty} E_n) = P(\bigcup_{i=1}^{\infty} G_i) \overset{\text{Axiom\#3}}{=\!=} \lim_{n\to\infty} \sum_{i=1}^{n} P(G_i) = \lim_{n\to\infty} P(\bigcup_{i=1}^{n} G_i)$
$= \lim_{n\to\infty} P(E_n)$ $\quad (G_n = E_n - E_{n-1}$, mutually exclusive$)$

$B_k = \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} A_n$ : decreasing, $C_k = \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} A_n$ : increasing

## Conditional Probability

$P(A \mid B) = \dfrac{P(A \cap B)}{P(B)}$ = conditional probability of event $A$ given $B$

Theorem (Reduction of Sample Space) $\Omega$: sample space, $P(B) > 0$

1. $P(A \mid B) \geq 0$ for any event $A$ $\quad$ 2. $P(\Omega \mid B) = 1$
3. $A_1, A_2 \dots$ are mutually exclusive, then $P(\bigcup_{i=1}^{\infty} A_i \mid B) = \sum_{i=1}^{\infty} P(A_i \mid B)$

Tool #1: Multiplication Rule

$P(\bigcap_{i=1}^{n} A_i) = P(A_1) P(A_2 \mid A_1) \dots P(A_n \mid A_1 \cap A_2 \dots A_{n-1})$

Tool #2: Total Probability Theorem

$A_1, \dots, A_n$ are partition of $\Omega$ and mutually exclusive, $P(A_i) > 0$

$P(B) = P(A_1 \cap B) + \dots + P(A_n \cap B) = P(A_1) P(B \mid A_1) + \dots + P(A_n) P(B \mid A_n)$

Tool #3: Bayes' Rule

$P(A_i \mid B) = \dfrac{P(A_i) P(B \mid A_i)}{P(B)} = \dfrac{P(A_i) P(B \mid A_i)}{P(A_1) P(B \mid A_1) + \dots + P(A_n) P(B \mid A_n)}$

## Independence

$P(A \cap B) = P(A) P(B)$, if $P(B) > 0$, then $P(A \mid B) = P(A)$

If $A, B$ are independent, then $A, B^c$ are also independent

(pf) $P(A \cap B^c) = P(A) - P(A \cap B) = P(A) - P(A) P(B) = P(A) \underbrace{(1 - P(B))}_{P(B^c)}$

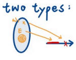
$A_1, \dots, A_n$ are independent if $P(\bigcap_{i \in S} A_i) = \prod_{i \in S} P(A_i)$ for every $S \subseteq \{1, \dots, n\}$

Conditionally independent if $P(\bigcap_{i \in S} A_i \mid B) = \prod_{i \in S} P(A_i \mid B)$

## Random Variable

A function that maps each outcome to a real number

Noted as " $X: S \to S_X$, $S_X = \{X(\omega) \mid \omega \in S . X \in \mathbb{R}\}$
two types: 1. discreate: take values over a discreate range
2. continuous: take values over a continuous range



*(range: $[0,1]$)

CDF of $X$: $F_X(t) = P(X \leq t)$, $\forall t \in \mathbb{R}$, $F_X(t) \in [0, 1]$
Cumulative Distribution Function

- $X \leq a$: $F_X(a)$ · $X > a$: $1 - F_X(a)$ · $X < a$: $F_X(a^-)$ or $\lim_{t \to a^-} F_X(t)$ or $F_X(a) - p(X=a)$
- $X \geq a$: $1 - F_X(a^-)$ · $X = a$: $F_X(a) - F_X(a^-)$

Property: 1. $F_X$ is non-decreasing 2. $\lim_{t\to\infty} F_X(t) = 1$ 3. $\lim_{t\to-\infty} F_X(t) = 0$.
4. $F_X$ is right continuous: $F_X(t^+) = F_X(t)$

$a < X \leq b$ $\quad F_X(b) - F_X(a)$
$a < X < b$ $\quad F_X(b^-) - F_X(a)$
$a \leq X \leq b$ $\quad F_X(b) - F_X(a^-)$
$a \leq X \leq b$ $\quad F_X(b^-) - F_X(a^-)$

PMF of $X$: $p(\cdot)$ of $X$ is the function that satisfies
Probability Mass Function

1. $P(x_i) = P(X = x_i)$ $\quad$ 2. $p(x) = 0$ if $x \notin \{x_1, x_2, x_3\}$ $\quad$ 3. $\sum_{i=1}^{\infty} p(x_i) = 1$

# Special Discrete Random Variables

## 1. Bernoulli Random Variables ( $X \sim \text{Bernoulli}(p)$ )

PMF : $P(X=k) = \begin{cases} p & \text{if } k=1 \\ 1-p & \text{if } k=0 \\ 0 & \text{otherwise} \end{cases}$  ( $X \sim \text{Binomial}(1,p)$ )

$E[X] = p$ , $Var[X] = p(1-p)$

1 experiment trial with 2 possible outcomes

## 2. Binomial Random Variables ( $X \sim \text{Binomial}(n,p)$ )

PMF : $P(X=k) = \begin{cases} C^n_k p^k (1-p)^{n-k} & \text{if } k=0,1,2,\dots,n \\ 0 & \text{otherwise} \end{cases}$

$E[X] = np$, $Var[X] = np(1-p)$

n repetitions of the same Bernoulli experiment

Q (for fixed k, under $p=?$ is $P(X=k)$ is max) : $P(X=k) = C^N_k p^k (1-p)^{N-k} = f(p)$

$\Rightarrow \ln f(p) = \ln C^N_k + k \cdot \ln p + (N-k) \cdot \ln(1-p)$

$\Rightarrow \frac{d}{dp} \ln f(p) = \frac{k}{p} - \frac{N-k}{1-p} = 0 \Leftrightarrow k(1-p) = p(N-k) \Leftrightarrow \boxed{p = \frac{k}{N}}$

## 3. Poisson Random Variables ( $X \sim \text{Poisson}(\underbrace{\lambda}_{\text{Average Rate}}, \overbrace{T}^{\text{Duration of observation window}})$ )

PMF : $P(X=n) = \frac{e^{-\lambda T} (\lambda T)^n}{n!}$ , $n = 0, 1, \dots$  ( $X \sim \text{Binomial}(n, \frac{\lambda}{n})$ )

$E[X] = Var[X] = \lambda T$

$\Rightarrow P(X=k) = C^n_k \cdot (\frac{\lambda}{n})^k \cdot (1-\frac{\lambda}{n})^{n-k} = \frac{n \cdot (n-1) \cdots (n-k+1)}{k!} (\frac{\lambda}{n})^k \frac{(1-\frac{\lambda}{n})^n}{(1-\frac{\lambda}{n})^k}$

$\left[ n \to \infty : \frac{n \cdot (n-1) \cdots (n-k+1)}{n^k} \to 1, \ (1-\frac{\lambda}{n})^n = \left((1-\frac{\lambda}{n})^{-\frac{n}{\lambda}}\right)^{-\lambda} = e^{-\lambda}, \ (1-\frac{\lambda}{n})^k \to 1 \right]$

as $n \to \infty$, $P(X=k) \to \frac{e^{-\lambda} \cdot \lambda^k}{k!}$

$P(\Omega) = \sum_{n=0}^{\infty} P(X=n) = e^{-\lambda T} \underbrace{\sum \frac{(\lambda T)^n}{n!}}_{e^{\lambda T}} = 1.$

Average rate is known and static

$\text{Poisson}(\lambda_1, T) + \text{Poisson}(\lambda_2, T) = \text{Poisson}(\lambda_1 + \lambda_2, T)$

## 4. Geometric Random Variables ( $X \sim \text{Geometric}(p)$ )

PMF : $P(X=k) = \begin{cases} (1-p)^{k-1} p & \text{if } k=1,2,3,\dots \\ 0 & \text{otherwise} \end{cases}$

Repetitions of the same Bernoulli experiment

$E[X] = \frac{1}{p}$ , $E[X^2] = \frac{2-p}{p^2}$ , $Var[X] = \frac{1-p}{p^2}$

$P(X = n+m \mid X > m) = P(X = n)$
$P(X > n+m \mid X > m) = P(X > n)$  ( Memoryless Property )

## 5. Discrete Uniform Random Variables

PMF : $P(X=k) = \frac{1}{b-a+1}$ , $k = a, a+1, \dots, b$

$E[X] = \frac{a+b}{2}$ . $Var[X] = \frac{(b-a+1)^2 - 1}{12}$

# Shannon Entropy

$H(X) := -\sum_i p(x_i) \log p(x_i) \equiv \mathbb{E}[-\log p(X)]$  base 2 : bits, base e : nats

# Expected Value ( $E[X] \equiv \mu_X$ )

$E[X] := \sum_{x \in S} x \cdot \underbrace{p_X(x)}_{\text{PMF}} = \sum_{i=1}^{\infty} (x_i - x_{i-1})(1 - F_X(x_i^-))$   $E[\alpha X + \beta] = \alpha E[X] + \beta$

$E[g(X)] := \sum_{x \in S} g(x)\, p_X(x)$ (LOTUS)   $E[g(X) + h(X)] = E[g(X)] + E[h(X)]$

# Moments

$E[X^n]$ : n-th moment   $E[e^{tX}]$ : moment generating function

$E[(X - \mu_X)^n]$ : n-th central moment

# Variance (2nd central moment)

$Var[X] := E[(X-\mu_X)^2] = \sum_{x \in S} (x - \mu_X)^2 \cdot p_X(x) = E[X^2] - (E[X])^2$

Property :

1. $Var[X+c] = Var[X]$
2. $Var[aX] = a^2 Var[X]$
3. $Var[|X|] \neq Var[X]$
4. $E[X^2] \ge (E[X])^2$

Riemann Rearrangement Theorem : let $\{a_n\}$ be a sequence of number if 1. $\sum_{n=1}^{\infty} a_n$ converges (收敛). 2. $\sum_{n=1}^{\infty} |a_n| = \infty$. then for any $B \in \mathbb{R} \cup \{\infty\}$ there exists a rearrangement $\{b_n\}$ of $\{a_n\}$ such that $\sum_{n=1}^{\infty} b_n = B$.

Existence of Moments : if $E[|X^n|] < \infty$ then $E[X^n]$ exists

· if $E[|X|^{n+1}] < \infty$, then $E[|X|^n] < \infty$

# Moment Generating Functions

$M_X(t) = E[e^{tX}]$ , $E[X^n] = M_X^{(n)}(0)$

Ex $X \sim N(\mu, \sigma^2)$ , $M_X(t) = e^{(\mu t + \frac{\sigma^2 t^2}{2})}$

$X \sim \text{Exp}(\lambda)$ , $M_X(t) = \frac{\lambda}{\lambda - t}$ for $t < \lambda$

Borel-Cantelli Lemma : if $\sum_{n=1}^{\infty} P(A_n) < \infty$, then $P(\bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} A_n) = 0$.

Borel Zero-one law : Let $A_1, A_2, \dots$ be a countably infinite sequence of events. Then, $P(\bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} A_n) = \begin{cases} 0, & \text{if } \sum_{n=1}^{\infty} P(A_n) < \infty \\ 1, & \text{if } \sum_{n=1}^{\infty} P(A_n) = \infty \end{cases}$

if $X_1, \dots, X_n$ are "independent geometry r.v. then $Y = \min(X_1, \dots, X_n)$ is also a geometry r.v. with $p = 1-(1-p)^n$

if $L(\alpha) = C^N_k \cdot \alpha^k \cdot (1-\alpha)^{N-k}$. which $\alpha$ can maximize $L(\alpha)$ is solved by calculate $L'(\alpha) = 0$.

# Probability of Winning if Switch Door

$P(E_{M,N}) = \frac{N-M}{N} \cdot \frac{\overset{\text{Empty}}{N-M-1}}{N-2} + \frac{M}{N} \cdot \frac{N-M}{N-2}$

$H(X) = -\sum_{i=1}^{n} p_i \ln p_i = \sum_{i=1}^{n} p_i \cdot \ln \frac{1}{p_i} = \sum_{i=1}^{n} \ln (\frac{1}{p_i})^{p_i} = \ln (\prod_{i=1}^{n} (\frac{1}{p_i})^{p_i})$

and by weighted inequality of arithmetic and geometric means, we have

$\frac{\frac{1}{p_1} \cdot p_1 + \frac{1}{p_2} \cdot p_2 + \dots + \frac{1}{p_n} \cdot p_n}{\sum_{i=1}^{n} p_i} \ge \left[ \prod_{i=1}^{n} (\frac{1}{p_i})^{p_i} \right]^{\frac{1}{\sum_{i=1}^{n} p_i}} \Rightarrow n \ge \prod_{i=1}^{n} (\frac{1}{p_i})^{p_i}$

$\because \ln$ is an increasing function $\therefore \ln n \ge \ln \prod_{i=1}^{n} (\frac{1}{p_i})^{p_i} = H(X)$

and we also know that equality holds if $(\frac{1}{p_1})^{p_1} = (\frac{1}{p_2})^{p_2} = \dots = (\frac{1}{p_n})^{p_n}$

$\Rightarrow p_1 = p_2 = \dots = p_n = \frac{1}{n}$. In conclusion, when the PMF $\{p_i\}_{i=1}^{n}$ : $p(x) = \begin{cases} 1/n, & X = 1, 2, 3, \dots, n \\ 0, & \text{otherwise} \end{cases}$

$H(X)$ has maximum value $\ln n$

# Continuous Random Variables

PDF of X : For $\underbrace{f_X(x)}_{\text{PDF}}$, $P(X \in B) = \int_B f_X(x) dx$. check vaild :

1. $P(X \in \mathbb{R}) = 1 \Leftrightarrow \int_{-\infty}^{\infty} f_X(x) dx$
2. $P(X \in A) \ge 0 \ \forall A \Leftrightarrow \int_A f_X(x) dx \ge 0 \ \forall A$
3. $P(X \in \bigcup_{i \ge 1} A) = \sum_{i \ge 1} P(X \in A_i) \Leftrightarrow \int_{\bigcup_{i \ge 1} A_i} f_X(x) dx = \sum_{i \ge 1} \int_{A_i} f_X(x) dx$

(don't need to check, $\because$ hold by the definition of integration)

CDF $\Leftrightarrow$ PDF : $F_X(t) = P(X \le t) = \int_{-\infty}^{t} f_X(x) dx$. $f_X(X) = F_X'(X)$

$\Rightarrow$ if $f_X(\cdot)$ is continuous at $X_0$, then $F_X'(X_0) = f_X(X_0)$

Property: 1. $f_X(X)$ could be $\ge 1$ because it doesn't have meaning as a single point 2. $p(a < X < b) = p(a \le X < b) = p(a < X \le b) = p(a \le X \le b) = \int_a^b f_X(x) dx$

$E[X] = \int_0^{\infty} (1 - F_X(t)) dt - \int_0^{\infty} \underbrace{F_X}_{\text{CDF}}(-t) dt$

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

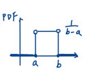
$E[g(X)] = \int_{-\infty}^{\infty} g(x) \cdot f_X(x) dx$

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

# Special Continuous Random Variables

## 1. Continuous Uniform Random Variables ( $X \sim \text{Unif}(a,b)$ )

PDF of X : $f_X(X) = \frac{1}{b-a} (a < X < b)$, 0 (otherwise)




CDF of X: $F_X(X \le t) = t$

· if we take CDF of X : $F_X(X)$ as r.v. then it's a kind of uniform r.v. : $F_X(X) \sim \text{Unif}(0,1)$

Inverse Transform Sampling (ITS) : given CDF: $F(t)$, generate r.v.

1. choose $U \sim \text{Unif}(0,1)$ 2. $X = F^{-1}(U)$, where $F^{-1}(u) := \inf\{z \mid : F(z) \ge u\}$

$P[F^{-1}(U) \le x] = P[F(F^{-1}(U)) \le F(x)] = P[U \le F(x)] = F(x)$

F(x) is non-decreasing   def. of unif

## 2. Standard Normal Random Variables ( $X \sim \mathcal{N}(0,1)$ )

PDF : $f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ , $E[X] = 0$ , $Var[X] = 1$

CDF : $\Phi(t) := P(X \le t) = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$ , $\Phi(\infty) = 1$ , $\Phi(0) = \frac{1}{2}$

### From Standard Normal to Normal

$Y = aX + b$  CDF : $F_Y(t) = P(Y \le t) = P(aX + b \le t) = \begin{cases} P(X \le \frac{t-b}{a}) & \text{if } a > 0 \quad F_X(\frac{t-b}{a}) \\ P(X \ge \frac{t-b}{a}) & \text{if } a < 0 \quad 1 - F_X(\frac{t-b}{a}) \end{cases}$

PDF : $\frac{d}{dt} F_Y(t) = \begin{cases} \frac{1}{a} f_X(\frac{t-b}{a}) & \text{if } a > 0 \\ \frac{1}{a} (-f_X(\frac{t-b}{a})) & \text{if } a < 0 \end{cases}$

Normal Random Variables : $f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\left(\frac{-(x-\mu)^2}{2\sigma^2}\right)}$   $X \sim \mathcal{N}(\mu, \sigma^2)$

## 3. Exponential Random Variables ( $X \sim \text{Exp}(\lambda)$ )

PDF : $f_X(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \ge 0 \\ 0 & \text{otherwise} \end{cases}$

Under a larger $\lambda$, X is more likely to be smaller

CDF : $t \le 0$ : $F_X(t) = 0$
$t > 0$ : $F_X(t) = P(X \le t) = \int_0^t f_X(x) dx = \int_0^t \lambda e^{-\lambda x} dx = 1 - e^{-\lambda t}$

$P(X > s+t \mid X > t) = \frac{P(X > s+t)}{P(X > t)} = \frac{\int_{s+t}^{\infty} f_X(x) dx}{\int_t^{\infty} f_X(x) dx} = \frac{e^{-\lambda(s+t)}}{e^{-\lambda t}} = e^{-\lambda s} = P(X > s)$

$F_X(t) = P(\{w : \min(X_1(w), X_2(w)) \le t\}) = 1 - P(X_1 > t) P(X_2 > t) = 1 - e^{-(\lambda_1 + \lambda_2) t}$

$X_1 \sim \text{Exp}(\lambda_1)$, $X_2 \sim \text{Exp}(\lambda_2)$, $X = \min(X_1, X_2) \sim \text{Exp}(\lambda_1 + \lambda_2)$

# Binomial Expansion

$(x+y)^n = \sum_{i=0}^{n} C^n_i x^{(n-i)} y^i$   $C^n_0 + \dots + C^n_n = 2^n$

$e^x = \lim_{n \to \infty} (1 + \frac{x}{n})^n = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} \dots$

$\Gamma(n) = \int_0^{\infty} u^{n-1} e^{-u} du$   $\Gamma(\frac{3}{2}) = \frac{\sqrt{\pi}}{2}$ , $\Gamma(\frac{1}{2}) = \sqrt{\pi}$

$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$   $\sum_{n=1}^{\infty} \frac{1}{n^4} = \frac{\pi^4}{90}$   $\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n$   $\ln(1+x) = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{x^n}{n}$