Chapter 1. Preliminaries.

1.1 probability.

$S \triangleq$ sample space
$E \triangleq$ event
$P(E)$. the probability of event E. which needs to satisfy.

* A1 $0 \leq P(E) \leq 1$
* A2 $P(S) = 1$
* A3 $E_1, E_2, \dots$, mutually exclusive, $P(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} P(E_i)$

Some consequences:

* 1.1.1. If $E \subset F$, then $P(E) \leq P(F)$.

proof. $P(\bigcup_{i=1}^{\infty} \phi) = \sum_{i=1}^{\infty} P(\phi) \Rightarrow P(\phi) = 0$

$P(F) = P(E \cup F/E \cup \phi \cdots) = P(E) + P(F/E). \geq P(E)$

#

* 1.1.2. $P(E^c) = 1 - P(E)$.

proof. $1 = P(S) = P(E \cup E^c) = P(E) + P(E^c)$. #

* 1.1.3 E_i mutually exclusive, $P(\bigcup_{i=1}^{n} E_i) = \sum_{i=1}^{n} P(E_i)$.

Obviously obtained from A3.

* 1.1.4. $P(\bigcup_{i=1}^{\infty} E_i) \leq \sum_{i=1}^{\infty} P(E_i)$. (Boole's inequality).

proof. we define $F_1 = E_1$, $F_n = E_n - (\bigcup_{i=1}^{n-1} E_i)$

F_i mutually exclusive, $\bigcup_{i=1}^{n} E_i = \bigcup_{i=1}^{n} F_i$, $F_n \subseteq E_n$

$P(\bigcup_{i=1}^{\infty} E_i) = P(\bigcup_{i=1}^{\infty} F_i) = \sum_{i=1}^{\infty} P(F_i) \leq \sum_{i=1}^{\infty} P(E_i)$.

#

Continuity.

increasing sequence: $E_n \subset E_{n+1}$, $n \geq 1$.
decreasing sequence: $E_n \supset E_{n+1}$, $n \geq 1$.
The limits: – increasing seq: $\lim_{n\to\infty} E_n = \bigcup_{n=1}^{\infty} E_n$.
– decreasing seq: $\lim_{n\to\infty} E_n = \bigcap_{n=1}^{\infty} E_n$.

* Prop 1.1.1 E_n increasing / decreasing seq., then

$$\lim_{n\to\infty} P(E_n) = P(\lim_{n\to\infty} E_n).$$

proof. ① increasing seq.

We define $F_1 = E_1$, $F_n = E_n - (\bigcup_{i=1}^{n-1} E_i) = E_n - E_{n-1}$.

F_i mutually exclusive. $\bigcup_{i=1}^{n} E_i = \bigcup_{i=1}^{n} F_i = E_n$

$$P(\lim_{n\to\infty} E_n) = P(\bigcup_{i=1}^{\infty} E_i) = P(\bigcup_{i=1}^{\infty} F_i) = \sum_{i=1}^{\infty} P(F_i)$$
$$= \lim_{n\to\infty} \sum_{i=1}^{n} P(F_i) = \lim_{n\to\infty} P(\bigcup_{i=1}^{n} F_i) = \lim_{n\to\infty} P(E_n)$$

② decreasing seq.

E_n^c increasing seq.

$$1 - \lim_{n\to\infty} P(E_n) = \lim_{n\to\infty} P(E_n^c) = P(\lim_{n\to\infty} E_n^c) = P((\bigcap_{n=1}^{\infty} E_n)^c)$$
$$= 1 - P(\lim_{n\to\infty} E_n)$$

$\Rightarrow \lim_{n\to\infty} P(E_n) = P(\lim_{n\to\infty} E_n)$.

#

* Prop 1.1.2 (Borel – Cantelli). If $\sum_{i=1}^{\infty} P(E_i) < \infty$, then $P(\text{an infinite number of } E_i \text{ occur}) = 0$

proof. [an infinite number of E_i occur] $= \bigcap_{n=1}^{\infty} \bigcup_{i=n}^{\infty} E_i$

$[\bigcup_{i=n}^{\infty} E_i]$ is a decreasing seq. for n.

Thus, $P(\text{an infinite number of } E_i \text{ occur})$

$$= P(\bigcap_{n=1}^{\infty} \bigcup_{i=n}^{\infty} E_i) = \lim_{n\to\infty} P(\bigcup_{i=n}^{\infty} E_i)$$

$$P(\bigcup_{i=n}^{\infty} E_i) \le \sum_{i=n}^{\infty} P(E_i) = 0$$

#

* Prop 1.1.3 (converse to B–C). If $E_1, E_2, \cdots$ independent, $\sum_{i=1}^{\infty} P(E_i) = \infty$, then $P(\text{an infinite number of } E_i \text{ occur}) = 1$.

proof. $P(\text{an infinite number of } E_i \text{ occur})$

$$= \lim_{n\to\infty} P(\bigcup_{i=n}^{\infty} E_i) = 1 - \lim_{n\to\infty} P(\bigcap_{i=n}^{\infty} E_i^c).$$
$$= 1 - \lim_{n\to\infty} \prod_{i=n}^{\infty} P(E_i^c)$$

$$\prod_{i=n}^{\infty} P(E_i^c) = \prod_{i=n}^{\infty} (1 - P(E_i)) \le \prod_{i=n}^{\infty} e^{-P(E_i)} = e^{-\sum_{i=n}^{\infty} P(E_i)} \to 0$$

#

Random Variable.

* distribution Function F for r.v. X.

$F(x) = P(X \le x)$. $\quad \bar{F}(x) = 1 - F(x)$.

if X discrete $\quad F(x) = \sum_{y \le x} P(X = y)$.

continuous. $\quad \exists$ density function $f(x)$.

$$P(x \in B) = \int_B f(x)\,dx.$$

* Joint distribution $\quad F(x,y) = P(X \le x, Y \le y)$.

and $\quad F_X(x) = P(X \le x) = \lim_{y \to \infty} F(x,y)$.

$\quad F_Y(y) = P(Y \le y) = \lim_{x \to \infty} F(x,y)$.

if X, Y independent. $\quad F(x,y) = F_X(x) \cdot F_Y(y)$.

Expected Value

* Expection $\quad E(X) = \int_{-\infty}^{+\infty} x\,dF(x)$.

$$= \begin{cases} \int_{-\infty}^{+\infty} x f(x)\,dx. & \text{continuous.} \\ \sum x\,p(x) & \text{discrete.} \end{cases}$$

if $h(x)$ is a r.v.

$$E\,h(x) = \int_{-\infty}^{+\infty} h(x)\,dF(x).$$

* Variance $\quad \mathrm{Var}\,X = E(X-EX)^2 = E(X^2 - 2X \cdot EX + (EX)^2)$
$\quad = EX^2 - (EX)^2$

* Covariance $\quad \mathrm{Cov}(X,Y) = E(X-EX)(Y-EY)$ (linear relations).
$\quad = E\,XY - EX \cdot EY$.

* Uncorrelated $\quad \Leftrightarrow \quad \mathrm{Cov}(X,Y) = 0$
* Independent $\quad F_X(x) F_Y(y) = F(x,y) \Rightarrow \mathrm{Cov}(X,Y) = 0$

$X \sim N(0,1)$. $Y = X^2$. X, Y is not independent

$\mathrm{Cov}(X,Y) = E\,XY - EX \cdot EY = E X^3 = 0$ (y轴对称).

* Some properties. $\quad E \sum_{i=1}^{n} X_i = \sum_{i=1}^{n} EX_i$

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

* Example 1.3 (a).

Example 1.3(a) ***The Matching Problem.*** At a party n people put their hats in the center of a room where the hats are mixed together Each person then randomly selects one We are interested in the mean and variance of X—the number that select their own hat

To solve, we use the representation

$$X = X_1 + X_2 + \cdots + X_n.$$

$$X_i = \begin{cases} 1 & i \text{ got his hat} \\ 0 & \text{otherwise} \end{cases}$$

$$EX_i = \frac{1}{n}, \quad \text{Var } X_i = \frac{1}{n} - \frac{1}{n^2}$$

$$EX = E\sum_{i=1}^n X_i = \sum_{i=1}^n EX_i = 1.$$

$$\text{Cov}(X_i, X_j) = EX_iX_j - EX_iEX_j.$$

$$X_iX_j = \begin{cases} 1 & i \text{ and } j \text{ got their hats} \\ 0 & \text{otherwise.} \end{cases}$$

$$\begin{aligned} EX_iX_j &= P(X_i = 1, X_j = 1) \\ &= P(X_i = 1 \mid X_j = 1)\, P(X_j = 1) \\ &= \frac{1}{n} \times \frac{1}{n-1} \end{aligned}$$

$$\text{Cov}(X_i, X_j) = \frac{1}{n}\frac{1}{n-1} - \frac{1}{n^2}$$

$$\begin{aligned} \text{Var } X = \text{Var} \sum_{i=1}^n X_i &= \sum_{i=1}^n \text{Var } X_i + 2\sum_{i<j} \text{Cov}(X_i, X_j) \\ &= 1 - \frac{1}{n} + 2\binom{n}{2}\left(\frac{1}{n}\frac{1}{n-1} - \frac{1}{n^2}\right) \\ &= 1 - \frac{1}{n} + 1 - \frac{n-1}{n} = 1. \end{aligned}$$

#

*Example 1.3(b) $A_1, A_2, \cdots, A_n$ some events.

$$I_j \begin{cases} 1 & A_j \text{ occur} \\ 0 & \text{otherwise} \end{cases}$$

$$N = \sum_{j=1}^n I_j.$$

$$(1-1)^N = \begin{cases} 1 & \text{if } N = 0 \\ 0 & \text{if } N > 0 \end{cases}$$

By the binomial theorem,

$$(1-1)^N = \sum_{i=0}^N \binom{N}{i}(-1)^i = \sum_{i=0}^n \binom{N}{i}(-1)^i$$

Denote $I = \begin{cases} 1 & \text{if } N > 0 \\ 0 & \text{if } N = 0 \end{cases}$

then $1 - I = \sum_{i=0}^n \binom{N}{i}(-1)^i \Rightarrow I = \sum_{i=1}^n \binom{N}{i}(-1)^{i+1}$

Take Expectation

$$\begin{aligned} EI &= P(N > 0) \\ &= P(\text{at least one of the } A_i \text{ occur}) \\ &= P\left(\bigcup_i^n A_i\right) \end{aligned}$$

$$EN = E\sum_{j=1}^n I_j = \sum_{j=1}^n P(A_j)$$

$$\begin{aligned} E\binom{N}{2} &= E[\text{the number of pairs of } A_j \text{ occur}] \\ &= E\sum_{i<j} I_iI_j \\ &= \sum_{i<j} P(A_iA_j) \end{aligned}$$

$$E\binom{N}{i} = \sum_{j_1<\cdots<j_i} P(A_{j_1}\cdots A_{j_i})$$

$$\Rightarrow P(\bigcup_i^n A_i) = \sum_{j=1}^n P(A_j) - \sum_{i<j} P(A_iA_j) + \cdots + (-1)^{n+1} P(A_1A_2\cdots A_n)$$

Denote $\quad I_r = \begin{cases} 1 & \text{if } N=r \\ 0 & \text{otherwise} \end{cases}$

$$I_r = \binom{N}{r}(1-1)^{N-r} = \binom{N}{r}\sum_{i=0}^{N-r}\binom{N-r}{i}(-1)^i$$

$$= \sum_{i=0}^{N-r}\binom{N}{r}\binom{N-r}{i}(-1)^i$$

$$= \sum_{i=0}^{N-r}\frac{N!}{(N-r)!\,r!}\frac{(N-r)!}{i!\,(N-r-i)!}(-1)^i = \sum_{i=0}^{n-r}\binom{N}{r+i}\binom{r+i}{r}(-1)^i$$

Take Expectation :

$$E\,I_r = \sum_{i=0}^{n-r}\binom{r+i}{r}(-1)^i E\binom{N}{r+i}$$

$$= \sum_{i=0}^{n-r}\binom{r+i}{r}(-1)^i \sum_{j_1<\cdots<j_{r+i}} P(A_{j_1}\cdots A_{j_{r+i}})$$

#

1.4 Moment Generating Function. Characterist Function and Laplace Transforms.

* MGF : $\quad \psi(t) = E(e^{t\bar{X}}) = \int e^{tx}\,dF(x).$

$\quad \psi'(t) = E\,\bar{X}e^{t\bar{X}} \quad . \quad \psi^{(n)}(0) = E\,\bar{X}^n$

$X \sim N(\mu_1, \sigma_1)$. $\quad Y \sim N(\mu_2, \sigma_2)$. independent

$$\psi_{X+Y}(t) = E(e^{t(X+Y)})$$

$$= E e^{tX}\,E e^{tY} = \psi_X(t)\,\psi_Y(t) = e^{(\mu_1+\mu_2)t + (\sigma_1^2+\sigma_2^2)\frac{t^2}{2}}$$

$\Rightarrow X+Y \sim N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$.

* CF : $\quad \phi(t) = E(e^{it\bar{X}})$

$\quad \psi(t_1, \cdots, t_n) = E\,e^{\sum_{j=1}^n t_j x_j}$

$\quad \phi(t_1, \cdots, t_n) = E\,e^{i\sum_{j=1}^n t_j x_j}$

* Example $\quad Z_1 \cdots, Z_n$ independent $\sim N(0,1)$

a_{ij}, μ_i constants.

$X_1 = a_{11}Z_1 + \cdots + a_{1n}Z_n + \mu$

$\vdots$

$X_m = a_{m1}Z_1 + \cdots + a_{mn}Z_n + \mu_m$

$\sum_{t=1}^m t_i X_i$ is a linear combination of the independent normal random variables,

$$E\left[\sum_{i=1}^m t_i X_i\right] = \sum_{i=1}^m t_i\mu_i$$

$$Var\left(\sum_{i=1}^m t_i X_i\right) = \sum_{i=1}^m\sum_{j=1}^m t_it_j\,Cov(X_i, X_j).$$

$E[e^Y] = \varphi_Y(t)|_{t=1} = e^{\mu+\frac{\sigma^2}{2}}$ $\quad (Y \sim N(\mu, \sigma^2))$

$\varphi(t_1, \cdots, t_m) = \exp\left[\sum_{i=1}^{m} t_i \mu_i + \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} t_i t_j \, Cov(X_i, X_j)\right]$

#

* Laplace transform

$$\tilde{F}(s) = \int_0^{+\infty} e^{-sx} \, dF(x).$$

s: real / complex number $a+bi$ with $a \geq 0$ (integral exist).

Laplace Transform uniquely determine the distribution

for any function $g(x)$. LT: $\tilde{g}(x) = \int_0^{+\infty} e^{-sx} \, dg(x)$

1.5 Conditional Expectation

If X, Y r.v. discrete. the conditional probability mass function of X given $Y=y$ is defined for all y such that $P(Y=y) > 0$

$$P(X=x \mid Y=y) = \frac{P(X=x, Y=y)}{P_Y(y)}$$

Conditional distribution function of X given $Y=y$

$$F(x|y) = P(X \leq x \mid Y=y).$$

Conditional expectation of X given $Y=y$

$$E(X|Y=y) = \int x \, dF(x|y) = \sum_x x \, p(X=x \mid Y=y).$$

If X, Y have joint probability density function $f(x,y)$. conditional probability density function of X given $Y=y$ is defined for all y s.t. $f_Y(y) > 0$ by

$$f(x|y) = \frac{f(x,y)}{f_Y(y)}$$

conditional probability distribution function of X given $Y=y$

$$F(x|y) = P(X \le x \mid Y=y) = \int_{-\infty}^{x} f(x|y)\,dx.$$

conditional expectation of X given $Y=y$

$$E(X|Y=y) = \int_{-\infty}^{+\infty} x f(x|y)\,dx$$

Define $E(X|Y)$ is a function Y.

$f(y) = E[X|Y=y]$. then $F(Y) = E[X|Y]$ is a r.v.

$$E[E[X|Y]] = EX = \int_{-\infty}^{+\infty} E[X|Y]\,d F_Y(y).$$

Y continuous, $EX = \int_{-\infty}^{+\infty} E[X|Y=y] f_Y(y)\,dy$

X, Y discrete. $EX = \sum_y E[X|Y=y] P(Y=y).$

$$= \sum_y \sum_x x P(X=x|Y=y) P(Y=y).$$

$$= \sum_x x P(X=x) = EX.$$

EX is weighted sum of $E[X|Y=y]$, weight $P(Y=y)$.

* Example 1.5 (A). $X_1, X_2, \cdots$ i.i.d r.v.

N is r.v. independent of $X_1, X_2, \cdots$

Compute moment generating function of $Y = \sum_{i=1}^{N} X_i$ conditional on N.

$$E[e^{t\sum_{i=1}^{N} X_i} \mid N=n] = E[e^{t\sum_{i=1}^{n} X_i}] = \prod_{i=1}^{n} E[e^{tX_i}] = (\psi_X(t))^n$$

$\psi_X(t)$ is $X_1, X_2, \cdots$ moment generating function

$\psi_X(t) = E[e^{tX}]$

$$\psi_Y = Ee^{t\sum_{i=1}^{N} X_i} = E[E(e^{t\sum_{i=1}^{N} X_i} \mid N=n)]$$

$$= E(\psi_X(t)^N)$$

$$EY = \psi_Y'(t)\big|_{t=0} = E[N \psi_X(t)^{N-1} \cdot \psi_X'(t)]\big|_{t=0}$$

$\psi_X(t)|_{t=0} = 1.$ $\psi_X'(t)|_{t=0} = EX.$

$EY = E[NEX] = EN\,EX$

$$EY^2 = \psi''_Y(t)|_{t=0} = E[N(N-1)\psi_X(t)^{N-2}(\psi'_X(t))^2 + N\psi_X(t)^{N-1}\psi''_X(t)]|_{t=0}$$

$$= E[N(N-1)(EX)^2] + E[N\,EX^2]$$

$$= EN^2(EX)^2 - EN(EX)^2 + EN\,EX^2$$

$$= EN^2(EX)^2 + EN\,\mathrm{Var}\,X.$$

$$\mathrm{Var}\,Y = EY^2 - (EY)^2 = EN^2(EX)^2 + EN\,\mathrm{Var}X - (EN)^2(EX)^2$$
$$= EN\,\mathrm{Var}\,X + (EX)^2\,\mathrm{Var}\,N.$$

* Example 1.5 (B).

If choose door 1: travel 2 hours to safety
door 2: travel 3 hours and come back
door 3: travel 5 hours and come back

X: # of hours take to reach safety

let Y: miner first time choose door Y

$$P(Y=1) = P(Y=2) = P(Y=3) = \frac{1}{3}$$

$$Ee^{tX} = E[E[e^{tX}|Y]] = \frac{1}{3}(E[e^{tX}|Y=1] + E[e^{tX}|Y=2] + E[e^{tX}|Y=3])$$

$$E[e^{tX}|Y=1] = e^{2t}$$

when $Y=2$, $X = 3+X'$, X': # of additional hours to safety after returning to the mine.

X' and X have the same distribution

$$E[e^{tX}|Y=2] = E[e^{3t+tX'}] = e^{3t}Ee^{tX}$$

$$E[e^{tX}|Y=3] = E[e^{5t+tX'}] = e^{5t}Ee^{tX}$$

$$\Rightarrow Ee^{tX} = \frac{1}{3}(e^{2t} + e^{3t}Ee^{tX} + e^{5t}Ee^{tX})$$

$$Ee^{tX} = \frac{e^{2t}}{3-e^{3t}-e^{5t}}$$

E is an event, $X = \begin{cases} 1 & \text{if } E \text{ occurs} \\ 0 & \text{if } E \text{ does not occur} \end{cases}$

$$P(E) = EX = E[E[X|Y]]$$

$$E[X|Y=y] = P(E|Y=y)$$

$$EX = \int P(E|Y=y)\,dF_Y(y).$$

* Example 1.5(C). n people n hat

pick a hat. if got the right hat go home. if not then pick again.

R_n: number of rounds for n people to get their hats

prove $E[R_n] = n$

Induction:

If $n=1$, $E[R_n] = 1$, obviously.

If $E[R_k] = k$ for $k = 1, 2, \cdots, n-1$

we denote M: # of matches in the first round.

$E[R_n] = \sum_{i=0}^{n} E[R_n | M=i]\, p(M=i)$

If $M=i$ $\quad E[R_n | M=i] = 1 + E[R_{n-i}]$

$$E[R_n] = \sum_{i=0}^{n} [1 + E[R_{n-i}]]\, p(M=i)$$
$$= \sum_{i=0}^{n} p(M=i) + E[R_n]\, p(M=0) + \sum_{i=1}^{n} E[R_{n-i}]\, p(M=i)$$
$$= 1 + E[R_n]\, p(M=0) + \sum_{i=1}^{n} (n-i)\, p(M=i)$$
$$= 1 + E[R_n] p(M=0) + n \sum_{i=1}^{n} p(M=i) - \sum_{i=1}^{n} i\, p(M=i)$$
$$= 1 + E[R_n] p(M=0) + n(1 - p(M=0)) - EM$$

$$E[R_n] = E[R_n]\, p(M=0) + n(1 - p(M=0))$$
$$\Rightarrow E[R_n] = n$$

* Example 1.5 (d). X, Y i.i.d. $\sim F, G$.

$$X+Y \triangleq F * G.$$

$$F*G(a) = P(X+Y \le a)$$
$$= \int_{-\infty}^{+\infty} P(X+Y \le a \mid Y=y)\, dG(y)$$
$$= \int_{-\infty}^{+\infty} P(X \le a-y)\, dG(y)$$
$$= \int_{-\infty}^{+\infty} F(a-y)\, dG(y).$$

we denote $F*F \triangleq F_2$, $F * F_{n-1} = F_n$ — n Fold convolution

F_n is the distribution of the sum of n i.i.d random variable with distribution F.

* Example 1.5 (E) A, B candidates. $A - n$ votes $\quad n > m$.

$B - m$ votes

$P(A \text{ always ahead}) = \frac{n-m}{n+m}$.

we denote $P_{n,m} \triangleq A$ always ahead.

$P_{n,m} = P(A \text{ ahead} \mid A \text{ receive last vote}) \cdot \frac{n}{n+m}$
$\quad + P(A \text{ ahead} \mid B \text{ receive last vote}) \cdot \frac{m}{n+m}$

A receive last vote A always ahead:
A − n-1 vote B − m vote. , A always ahead : $P_{n-1,m}$.

B receive last vote. A always ahead:
A − n vote B − m-1 vote , A always ahead : $P_{n,m-1}$

$\Rightarrow P_{n,m} = \frac{n}{n+m} P_{n-1,m} + \frac{m}{n+m} P_{n,m-1}$

Induction on $n+m$. if $n+m=1$. $P_{1,0} = 1 = \frac{1}{1+0}$

assume $n+m=k$ is true

consider $n+m=k+1$. we have.

$$P_{n,m} = \frac{n}{n+m} P_{n-1,m} + \frac{m}{m+n} P_{n,m-1}$$

$$= \frac{n}{n+m} \times \frac{n-m-1}{n+m-1} + \frac{m}{m+n} \times \frac{n-m+1}{n+m-1}$$

$$= \frac{n^2-nm-n+mn-m^2+m}{(n+m)(n+m-1)} = \frac{(n-m)(n+m-1)}{(n+m)(n+m-1)} = \frac{n-m}{n+m}$$

1.6. Exponential distribution, memoryless, hazard rate function

* Exponential distribution.

continuous r.v. $X \sim$ exponential distribution with parameter λ, $\lambda > 0$, then its probability density function

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

cumulative distribute function:

$$F(x) = \begin{cases} 1 - e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

moment generating function

$$\varphi_X(t) = E[e^{tX}] = \int_0^{+\infty} e^{tx} \cdot \lambda e^{-\lambda x} \, dx$$

$$= \lambda \int_0^{+\infty} e^{(t-\lambda)x}\,dx$$

$$= \frac{\lambda}{t-\lambda} e^{(t-\lambda)x}\Big|_0^{+\infty} = \frac{\lambda}{\lambda-t} \quad (t<\lambda)$$

Expectation and Variance.

$$EX = \psi_X'(0) = \frac{\lambda}{(\lambda-t)^2}\Big|_{t=0} = \frac{1}{\lambda}$$

$$EX^2 = \psi_X''(0) = \frac{2\lambda}{(\lambda-t)^3}\Big|_{t=0} = \frac{2}{\lambda^2}$$

$$\text{Var } X = EX^2 - (EX)^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

* Memoryless.

$$P(X>s+t \mid X>t) = P(X>s). \qquad (s,t\geq 0).$$

$$\Leftrightarrow \frac{P(X>s+t,\, X>t)}{P(X>t)} = P(X>s)$$

$$\Leftrightarrow P(X>s+t) = P(X>s)\,P(X>t)$$

$$\Leftrightarrow \bar{F}(s+t) = \bar{F}(s)\,\bar{F}(t) \quad \Leftrightarrow \quad \bar{F}(x) = e^{-\lambda x} \quad (x>0)$$

Example 1.6 (A).

service time $\sim \exp(\lambda)$.

We suppose B complete his service and A start being serviced.

Because of memoryless, A and C have the same distribution at the time when A starts his service.

Hence, $P(A \text{ leave after } C) = \frac{1}{2}$ by symmetry.

Prob(A is the last one who leaves the office) $= \frac{1}{2}$

Example 1.6 (B).

$X_1, X_2, \cdots,$ i.i.d continuous c.d.f F

A record happens at n has value X_n if $X_n > \max\{X_1, \cdots X_{n-1}\}$

τ_i denote the time between i and ith record

Qestion $\tau_i \sim$?

$X_1, \cdots, X_n$ record time is identical to $F(X_1) \cdots, F(X_n)$.

$F(X_1) \cdots, F(X_n) \sim U(0,1) \Rightarrow \tau_i$'s distribution is independent of actual distribution F. Now we let F be $\exp(1)$

$R_i \triangleq$ ith record value.

$R_1 = X_1 \sim \exp(1)$. $\qquad R_2 \sim \exp(1) \mid X_2 > X_1$

Because of Memoryless R_2 has the same distribution as $x_1 +$ ind $\exp(1)$.

$R_2 \sim x_1 + x_2$, $x_1, x_2 \sim$ ind $\exp(1)$.

$R_i \sim x_1 + \cdots + x_i$, $x_1, \cdots, x_i \sim$ ind $\exp(1)$

the density $f_{R_i}(t) = \frac{e^{-t} t^{i-1}}{(i-1)!}$ (Gamma distribution)

$$p(\tau_i > k) = \int_0^{+\infty} P(\tau_i > k \mid R_i = t) \frac{e^{-t} t^{i-1}}{(i-1)!} dt$$

$$= \int_0^{+\infty} (1-e^{-t})^k e^{-t} \frac{t^{i-1}}{(i-1)!} dt$$

$\exp(\lambda) \Leftrightarrow \bar{F}(s+t) = \bar{F}(s)\bar{F}(t) \Leftrightarrow \bar{F}$ 是满足 $g(s+t) = g(s)g(t)$ 的 Function

$\Leftrightarrow g(x) = e^{-\lambda x}$

proof. We suppose $g(x)$ right continuous.

$g(\frac{2}{n}) = [g(\frac{1}{n})]^2$

$g(\frac{m}{n}) = [g(\frac{1}{n})]^m$

$g(1) = g(\frac{1}{n} \times n) = [g(\frac{1}{n})]^n$

$\Rightarrow g(\frac{m}{n}) = [g(1)]^{\frac{m}{n}}$, for all n, m.

由于右连续性: $\Rightarrow g(x) = [g(1)]^x$

而 $g(1) = [g(\frac{1}{2})]^2 > 0$

$\Rightarrow g(x) = e^{-\lambda x}$, 其中 $\lambda = -\log[g(1)]$

#

* failure rate / harzard rate.

X continuous, distribute function F

density function f.

$\lambda(t) = \frac{f(t)}{\bar{F}(t)}$ —— harzard rate.

$$P(X \in (t, t+dt) \mid X > t) = \frac{P(X \in (t,t+dt), X > t)}{p(X > t)}$$

$$= \frac{P(X \in (t, t+dt))}{p(X > t)} \xrightarrow{dt \to 0} \frac{f(t)}{\bar{F}(t)} = \lambda(t).$$

$\lambda(t)$ 是在 t 时刻坏掉的速率.

For exponential disthibution:

$\lambda(t) = \frac{\lambda e^{-\lambda t}}{e^{-\lambda t}} = \lambda$ independent of t

$\Rightarrow$ exp memoryless

$\lambda(t)$ uniquely determine r.v.

$\lambda(t) = \frac{f(t)}{\bar{F}(t)} = \frac{-\frac{d}{dt}\bar{F}(t)}{\bar{F}(t)}$

$$\Rightarrow -\int_0^t \lambda(t)\,dt + K = \int_0^t \frac{\frac{d}{dt}\bar{F}(t)}{\bar{F}(t)}\,dt = \int_0^t \frac{1}{\bar{F}(t)}\,d\bar{F}(t)$$
$$= \log \bar{F}(t)$$

$$\Rightarrow \bar{F}(t) = c\,e^{-\int_0^t \lambda(t)dt}$$

when $t=0 \Rightarrow c=1$

$$\Rightarrow \bar{F}(t) = e^{-\int_0^t \lambda(t)\,dt}$$

1.7 Probability Inequality

* Markov inequality.

Lem 1.7.1. $X \geq 0$, $\forall\, a > 0$

$$P(X \geq a) \leq \frac{EX}{a}$$

proof. let $I_{(X\geq a)} = 1$ if $X \geq a$

$I_{(X\geq a)} = 0$ if $X < a$.

Since $X > 0$, we have

$$a\, I_{(X\geq a)} \leq X$$

Then $EX \geq Ea\, E\, I_{(X\geq a)} = a \cdot P(X \geq a)$

$$\Rightarrow P(X \geq a) \leq \frac{EX}{a}.$$

#

Prep 1.7.2 chernoff bounds.

let X r.v. moment generating function $M(t) = E\,e^{tX}$

Then $a > 0$

$P(X \geq a) \leq e^{-at} M(t)$ $\quad \forall\, t > 0$

$P(X \leq a) \leq e^{-at} M(t)$ $\quad \forall\, t < 0$

proof. $t > 0$

$$p(X \geq a) = p(e^{tX} \geq e^{ta}) \leq \frac{E[e^{tX}]}{e^{ta}} = e^{-at} M(t).$$

$t < 0$

$$p(X \leq a) = p(e^{tX} \geq e^{ta}) \leq e^{-at} M(t).$$

#

Example 1.7(A) (Chernoff bounds for Possion r.v.)

$X \sim$ Possion (λ) $\quad M(t) = e^{\lambda(e^t - 1)}$

Chernoff Bounds for

$P(X \geq j)$ is

$$P(X \geq j) \leq e^{-jt} e^{\lambda(e^t - 1)}$$

minimize the bound. since $\lambda(e^t - 1) - jt$ convex.

take derivate: $\lambda e^t - j = 0 \Rightarrow e^t = \frac{j}{\lambda}$ provided $\frac{j}{\lambda} > 1$.

$$P(X \geq j) \leq e^{\lambda(\frac{j}{\lambda} - 1)} \cdot \left(\frac{j}{\lambda}\right)^{-j}$$

$= e^{-\lambda} \frac{(e\lambda)^j}{j^j}$

Prep 1.7.3 Jensen's Inequality.

f convex $Ef(Z) \geq f(EZ)$ provided that the Expectation exists.

proof. f has the Taylor's expansion.

Expand f at $\mu = EX$.

$$f(x) = f(\mu) + f'(\mu)(x-\mu) + \frac{1}{2} f''(\xi)(x-\xi)^2$$
$$\geq f(\mu) + f'(\mu)(x-\mu)$$
$$Ef(x) \geq f(E(x)).$$

#

1.8 Limit Theorem.

* Strong Law of Large Numbers

$X_1, X_2, \cdots,$ i.i.d mean μ. then

$$P\left(\lim_{n\to\infty} \frac{X_1+\cdots+X_n}{n} = \mu\right) = 1$$

$$\Rightarrow \lim_{n\to\infty} \frac{X_1+\cdots+X_n}{n} = \mu \quad \text{a.s. } 1.$$

* Central Limit Theorem.

$X_1, \cdots, X_n$ i.i.d. mean μ. Variance σ^2

$$\lim_{n\to\infty} P\left(\frac{\sum X_i - n\mu}{\sqrt{n}\,\sigma} \leq a\right) = \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx.$$

1.9. Stochastic Process.

$X : \{X(t), t \in T\}$ collection of r.v. $\forall t \in T$, $X(t)$ is r.v.

if T: countable X – discrete – time S.P.

if T: continuous X – continuous – time S.P.

Realization of X : sample path

X has independent increments : if.

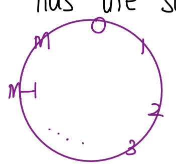
for $t_0 < t_1 < \cdots < t_n$

$X(t_1) - X(t_0)$, $X(t_2) - X(t_1)$, $\cdots$, $X(t_n) - X(t_{n-1})$ independent

X is stationary : if

$X(t+s) - X(t)$ has the same distribution for all t

Example 1.9 (A)



X_n is the position of the partical after its nth step.

$p(X_{n+1} = i+1 \mid X_n = i) = p(X_{n+1} = i-1 \mid X_n = i) = \frac{1}{2}$

where $i+1 = 0$ when $i = m$, and $i-1 = m$ where $i = 0$

Suppose that the particle starts at 0 and continuous to move around according to the above rule until all the nodes have been visited.

prob of i is the last visited ?