Probability & Statistics Notes

by Stone Sun

Probability : foundational subject

↳ unintuitive, often have bad intuition

How to study : no intuition

just apply logical framework

Mathematical framework : built around sets

A sample space : Ω, discrete or continous

An event : A, B, etc.

Probability : a function, denoted $P(A)$

Example :

1. Sample Space : $\Omega_1 = \{x \mid x \geq 0\}$

$\Omega_2 = \{x \mid x = 0, 1, 2, \cdots\}$

2. Event : $A = \{x \mid 0 \leq x \leq 600\}$

$B = \{3\}$

Probability axioms

Axiom 1 $\forall A,\ P(A) \geq 0$

Axiom 2 $P(\Omega) = 1$

Axiom 3 $A_1, A_2, \cdots$ are disjoint events,

then $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$

Def. If $A = \bigcup_{i=1}^{\infty} A_i$, $A_1, A_2, \cdots$ are disjoint, then

$A_1, A_2, \cdots$ is said to be a partition of A.

Frequency Interpretation

Results from axioms:

1. $P(A^c) = 1 - P(A)$
2. If $A \subset B$, then $P(B \cap A^c) = P(B) - P(A)$
3. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

naive definition (equally likely outcomes)

n_A: the number of sample points in A

N: the number of sample points in Ω

the probability of A: $P(A) = \frac{n_A}{N}$

Counting

(1) multiplication rule: $n_1 \times n_2 \times n_3 \times \cdots \times n_k$

another notation: $\prod_{i=1}^{k} n_i$

---

(2) permutations: $P_r^n = n(n-1)\cdots(n-r+1)$

Factorials: $n! = n(n-1)(n-2)\cdots 2 \times 1$

also define: $0! = 1$

or product notation: $n! \prod_{i=0}^{n-1}(n-i)$

Then $P_r^n = \frac{n!}{(n-r)!}$

(3) Combinations

choosing without replacement:

$$\binom{n}{r} = \frac{n!}{r!(n-r)!}$$

$$P_r^n = C_r^n \cdot r!$$

Example $n=8$, $r=3$:

$$C_3^8 = \binom{8}{3} = \frac{8!}{3!(8-3)!} = \frac{8\times 7\times 6\cdot 5!}{3\times 2\times 1\cdot 5!} = 56$$

Counting the complement:

at least : all − not at least

(4) without replacement without ordering

Stars and bars representation:

two possible case:

$\times\times||||$ $\qquad\qquad$ $\times||||\times$

$$\frac{6!}{2!\,4!}$$

r objects, n distinct containers

$\underbrace{\times\times\times\cdots\times}_{r}$ $\quad$ $\underbrace{|\,|\,|\cdots|}_{n-1}$ $\quad$ n-1 divisions

There are $r+n-1$ objects to arrange:

$$\frac{(r+n-1)!}{r!\,(n-1)!} = \binom{n+r-1}{r}$$

(3) Partitioning

Example: $n = \frac{10!}{3!\,3!\,4!} = 4200$

Another way: $n = \binom{10}{3}\binom{7}{3}\binom{4}{4}$

n distinct objects, k distinct groups:

$$N = \binom{n}{n_1, n_2, \cdots n_k} = \frac{n!}{n_1!\, n_2! \cdots n_k!}$$

(multinomial coefficient)

Summary

	Without Replacement	With Replacement
Ordered	$\frac{n!}{(n-r)!}$	n^r
Unordered	$\binom{n}{r}$	$\binom{n+r-1}{r}$

Partitioning: $N = \frac{n!}{n_1!\, n_2! \cdots n_k!}$

General strategies:

(1) multiplication principle

(2) over counting and correcting

(3) isomorphic problems

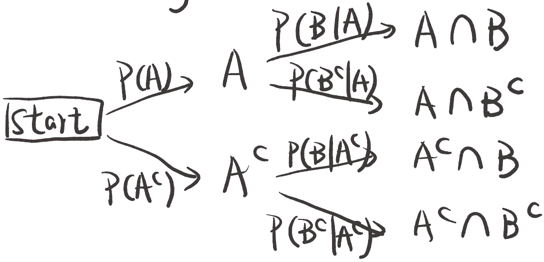
Conditional Probability

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

$$P(A\cap B) = P(B|A)\,P(A) = P(A|B)\,P(B)$$

$$P(A\cap B\cap C) = P((A\cap B)\cap C) = P(C|A\cap B)\,P(A\cap B)$$
$$= P(C|A\cap B)\,P(B|A)\,P(A)$$

Tree Diagram



Independence

Independent means anyone of the following holds:

$$P(B|A) = P(B)$$

$$P(A|B) = P(A)$$

$$P(A\cap B) = P(A)\,P(B)$$

otherwise, A & B are dependent.

Three events:

$P(A\cap B) = P(A)\,P(B)$, $P(A\cap C) = P(A)\,P(C)$

$P(B\cap C) = P(B)\,P(C)$, $P(A\cap B\cap C) = P(A)\,P(B)\,P(C)$

Law of total probability

$$P(B) = P(B \cap A) + P(B \cap A^c)$$
$$= P(B|A)P(A) + P(B|A^c)P(A^c)$$

For A_i, $i = 1, \cdots, n$:

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

Bayes' Rule

$$P(A) > 0, \quad P(B_j|A) = \frac{P(A|B_j)\,P(B_j)}{P(A)}$$

$$P(A) = \sum_{j=1}^{n} P(A|B_i)\,P(B_i)$$

General form: $$P(B_j|A) = \frac{P(A|B_j)\,P(B_j)}{\sum_{j=1}^{n} P(A|B_j)\,P(B_i)}$$

Probability of odds:

$$\text{Odds}(A) = \frac{P(A)}{P(A^c)} = \frac{P(A)}{1 - P(A)}$$

Bayes' rule for odds:

$$\underbrace{\frac{P(B|A)}{P(B^c|A)}}_{\text{posterior odds}} = \underbrace{\frac{P(A|B)}{P(A|B^c)}}_{\text{likelihood ratio}} \cdot \underbrace{\frac{P(B)}{P(B^c)}}_{\text{prior odds}}$$

2. Discrete Random Variables

Consider function $Y(\omega): \Omega \to \mathbb{R}$

Y is called a random variable.

Uppercase letters represent random variables,
Lowercase letters represent a particular value.

- countably infinite/finite: discrete
- countinous

The state space: the set $\Omega_Y \subseteq \mathbb{R}$ of all possible values $Y(\omega)$ can take.

Probability mass function

definition: $f_Y: \Omega_Y \to [0,1]$, $f_Y(y_i) = P(Y=y_i)$

properties (check valid): $0 \leq f(y_i) \leq 1$ for all i

$$\sum_{y_i \in \Omega_Y} f(y_i) = 1$$

cumulative distribution function

definition: $F_Y : \mathbb{R} \to [0,1]$

$$F_Y(y) = P(Y \le y)$$

Expected value: $E[Y] = \sum_{y_i \in \Omega_Y} y_i f(y_i)$

$$[E(Y)]^2 \neq E(Y^2)$$

moments: $E(Y^k)$ (kth moment)

$\sigma_k = E[(Y-m_1)^k]$ (kth central moment)

leptokurtic: "too peaked" to be normal

platykurtic: "too flat"

Variance: $Var(Y) = E(Y^2) - [E(Y)]^2$

standard deviation: $|\sqrt{Var(Y)}|$, denoted σ

properties: $E(a) = a$, $E(aY) = aE(Y)$

$$E(X+Y) = E(X) + E(Y)$$

$$E(c \cdot g(Y)) = c \cdot E(g(Y))$$

$$E\left(\sum_i c_i g_i(Y)\right) = \sum_i c_i E(g_i(Y))$$

Bonus: Birthday Problem

Assumptions

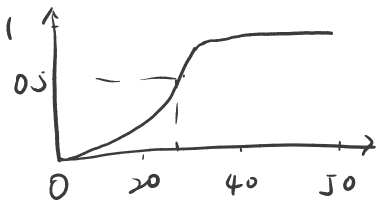
- 365 days in a year
- Birthdays are distributed every day
- Birthdays are independent

365^k

For complement: $P_K^{365} = 365 \cdot 364 \cdots (365-k+1)$

the probability:

$$1 - \frac{365 \cdot 364 \cdots (365-k+1)}{365^k}$$



(1) Bernouli distribution

y	0	1
$P(Y=y)$	$1-p$	p

Y with this pmf is said to have a Bernoulli Distribution

$E(Y) = p$, $Var(Y) = p(1-p)$

indicator : $I_A = \begin{cases} 1, & \text{if } A \text{ occurs} \\ 0, & \text{if } A \text{ doesn't occur} \end{cases}$

(2) Geometric Distribution

pmf of Y: $f(y) = (1-p)^{y-1} \cdot p$, $y = 1, 2, \cdots$

$Y \sim Geo(p)$

moments : $E(Y) = \frac{1}{p}$, $Var(Y) = \frac{1-p}{p^2}$

Proof. $E(Y) = \sum y \cdot P(Y=y)$

$$= p \sum \frac{d}{dq} q^y = p \cdot \frac{d}{dq} \cdot \sum q^y = p \cdot \frac{1}{(1-q)^2}$$

$(q = 1-p)$

$$Var(Y) = E(Y^2) - E(Y)^2$$
$$= E(Y^2) - E(Y) + E(Y) - E(Y)^2$$
$$= E(Y^2 - Y) + E(Y) - E(Y)^2$$

while $E(Y^2 - Y) = (1-p)p \sum y(y-1)(1-p)^{y-2}$

$$= (1-p) \cdot p \cdot \frac{d^2}{dq^2} \sum q^y$$
$$= (1-p) \cdot p \cdot \frac{2}{(1-q)^3} = \frac{2(1-p)}{p^2}$$

(3) Binomial Distribution

$$f(y) = \binom{n}{y} p^y (1-p)^{n-y}, \; y = 0, 1, \cdots, n$$

$Y \sim Bin(n, p)$

$E(Y) = np$, $Var(Y) = np(1-p)$

(4) Hypergeometric Distribution

population : N successes : r failures : $N-r$

pmf : $$f(y) = \frac{\binom{r}{y}\binom{N-r}{n-y}}{\binom{N}{n}}$$

* without replacement Y : the number of successes

$Y \sim hyper(N, n, r)$

$$E(Y) = \frac{rn}{N}$$

$$Var(Y) = n\left(\frac{r}{N}\right)\left(\frac{N-r}{N}\right)\left(\frac{N-n}{N-1}\right)$$

approximate : the binomial, $p = \frac{r}{N}$

(5) Poisson Distribution

n approaches infity : $np = \lambda$

$$P(Y=y) = \lim_{n\to\infty} \binom{n}{y}\left(\frac{\lambda}{n}\right)^y\left(1-\frac{\lambda}{n}\right)^{n-y} = \frac{e^{-\lambda}\lambda^y}{y!}$$

$$\lim_{n\to\infty} \binom{n}{y}\left(\frac{\lambda}{n}\right)^y\left(1-\frac{\lambda}{n}\right)^{n-y} = \lim_{n\to\infty} \frac{n(n-1)\cdots(n-y+1)}{y!}\left(\frac{\lambda}{n}\right)^y\left(1-\frac{\lambda}{n}\right)^{n-y}$$

$$= \lim_{n\to\infty} \frac{\lambda^y}{y!}\left(1-\frac{\lambda}{n}\right)^n \frac{n(n-1)\cdots(n-y+1)}{n^y}\cdot\left(1-\frac{\lambda}{n}\right)^{-y}$$

$$= \frac{\lambda^y}{y!}\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-y}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{y-1}{n}\right) = \frac{\lambda^y}{y!}\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n$$

$$= \frac{e^{-\lambda}\lambda^y}{y!}$$

$E(Y) = \lambda$, $Var(Y) = \lambda$

$$E(Y) = \sum_{y=0}^{\infty} y \cdot P(Y=y) = \sum_{y=0}^{\infty} y \cdot \frac{e^{-\lambda}\lambda^{y}}{y!}$$

$$= \lambda \cdot \sum_{y=1}^{\infty} \frac{e^{-\lambda} \cdot \lambda^{y-1}}{(y-1)!} = \lambda e^{-\lambda} \sum_{\gamma=0}^{\infty} \frac{\lambda^{\gamma}}{\gamma!} = \lambda e^{-\lambda} \cdot e^{\lambda}$$

$$= \lambda$$

$$Var(Y) = E(Y(Y-1)) + E(Y) - E(Y)^2$$

$$E(Y(Y-1)) = \sum_{y=2}^{\infty} y(y-1) \cdot \frac{e^{-\lambda}\lambda^{y}}{y!} = \lambda^2 \cdot \sum_{y=2}^{\infty} \frac{e^{-\lambda} \cdot \lambda^{y-2}}{(y-2)!}$$

$$= \lambda^2 e^{-\lambda} \sum_{\gamma=0}^{\infty} \frac{\lambda^{\gamma}}{\gamma!} = \lambda^2 e^{-\lambda} \cdot e^{\lambda} = \lambda^2$$

$$Var(Y) = \lambda^2 + \lambda - \lambda^2 = \lambda$$

2.3 Bounding probabilities

(1) Tail Sum formula

$$E(Y) = \sum_{i=1}^{n} P(Y \geq i)$$

Proof: $E(Y) = \sum_{k=0}^{n} k \cdot P(X=k)$

$= 0 \cdot P(X=0) + \cdots + n P(X=n)$

$= P(X=1) +$

$P(X=2) + P(X=2) +$

$\cdots +$

$P(X=n) + \cdots + P(X=n)$

$= P(X \geq 1) + P(X \geq 2) + \cdots + P(X \geq n)$

$= \sum_{k=1}^{n} P(X \geq k)$

(2) Markov's Inequality

Assume that the random variable is non-negative:

$$P(Y \geq a) \leq \frac{E(Y)}{a}$$

Proof. $E(Y) = \sum_{y=0}^{n} y \cdot P(Y=y) = \sum_{y=0}^{a-1} y P(Y=y) + \sum_{y=a}^{n} y P(Y=y)$

$\geq \sum_{y=a}^{n} y P(Y=y) \geq \sum_{y=a}^{n} a \cdot P(Y=y) = a \cdot P(Y \geq a)$

That is, $P(Y \geq a) \leq \frac{E(Y)}{a}$

(3) Chebyshev's Inequality

random variable Y, $E(Y) = \mu$, $Var(Y) = \sigma^2$

For any constant $k > 0$, $P(|Y-\mu| < k\sigma) \geq 1 - \frac{1}{k^2}$

$P(|Y-\mu| \geq k\sigma) \leq \frac{1}{k^2}$

Proof. $P(|Y-\mu|\geqslant k\sigma)=P((Y-\mu)^2\geqslant k^2\sigma^2)$

Let $X=(Y-\mu)^2$, $a=k^2\sigma^2$, then $P(X\geqslant a)\leqslant\frac{E(X)}{a}$

substitute the parameters, $P(X\geqslant a)\leqslant\frac{\sigma^2}{k^2\sigma^2}=\frac{1}{k^2}$

2.4 Moment Generating Functions

$$m(t) := E(e^{tY})$$

$$m_y(t) = E(e^{tY}) = \sum_{y=0}^{\infty} e^{ty} \cdot \frac{e^{-\lambda} \cdot \lambda^y}{y!} = e^{-\lambda} \cdot \sum_{y=0}^{\infty} \frac{(\lambda e^t)^y}{y!} = e^{-\lambda} \cdot e^{\lambda e^t}$$

$$= e^{\lambda(e^t - 1)}$$

$$\forall k > 0,\ m^{(k)}(0) = \left.\frac{d^k m(t)}{dt^k}\right|_{t=0} = \left.\frac{d^k E(e^{tY})}{dt^k}\right|_{t=0} = E(Y^k)$$

Proof.
$$\left.\frac{d^k m(t)}{dt^k}\right|_{t=0} = \left.\left(\frac{d^k}{dt^k} \sum_{y_i \in \Omega_Y} e^{ty_i} f(y_i)\right)\right|_{t=0} = \sum_i \left.\frac{d^k e^{ti}}{dt^k} f(y_i)\right|_{t=0}$$

$$= \sum_{y_i \in \Omega_Y} i^k e^{ti} f(y_i)\Big|_{t=0} = \sum_i i^k f(y_i) = E(Y^k)$$

Example. $Y \sim \text{Poisson}(\lambda)$. show $E(Y) = \lambda$, $Var(Y) = \lambda$.

$$m(t) = e^{\lambda(e^t - 1)}$$

$$\frac{d}{dt} m(t) = \lambda e^t \cdot e^{\lambda(e^t - 1)}\Big|_{t=0} = \lambda e^0 \cdot e^{\lambda(e^0 - 1)} = \lambda$$

$$E(Y) = \lambda$$

$$\frac{d^2}{dt^2} m(t) = \lambda e^t e^{\lambda(e^t - 1)} + (\lambda e^t)^2 e^{\lambda(e^t - 1)}\Big|_{t=0} = \lambda + \lambda^2$$

$$E(Y^2) = \lambda + \lambda^2$$

$$Var(Y) = E(Y^2) - E(Y)^2 = \lambda + \lambda^2 - \lambda^2 = \lambda$$

Probability generating functions

$$P(t) = E(t^Y) = P(Y=0) + t\,P(Y=1) + \cdots$$

Example $Y \sim Geo(p)$

$$P(t) = E(t^Y) = \sum_{y=1}^{\infty} t^y (1-p)^{y-1} p = pt \sum_{y=1}^{\infty} t^{y-1}(1-p)^{y-1}$$

$$= \frac{pt}{1-t(1-p)}$$

Properties: $P(1) = 1$, $P(0) = P(Y=0)$

$$\frac{d^k}{dt^k} P(t)\Big|_{t=0} = k!\,P(Y=k),$$

$$\frac{d^k}{dt^k} P(t)\Big|_{t=1} = E[Y(Y-1)\cdots(Y-k+1)]$$

Example $P(t) = e^{\lambda t} e^{-\lambda}$, for Poisson Distribution

Proof. $P(t) = E(t^Y) = \sum_{y=0}^{\infty} \frac{t^y e^{-\lambda} \lambda^y}{y!} = e^{-\lambda} \cdot \sum_{y=0}^{\infty} \frac{(t\lambda)^y}{y!}$

$$= e^{-\lambda} \cdot e^{\lambda t}$$

$$\frac{dP(t)}{dt} = e^{-\lambda} \cdot \lambda \cdot e^{\lambda t}\Big|_{t=1} = \lambda$$

$$\frac{d^2 P(t)}{dt^2} = e^{-\lambda} \lambda^2 e^{\lambda t}\Big|_{t=1} = \lambda^2 = E[Y(Y-1)]$$

$$Var(Y) = E[Y(Y-1)] + E(Y) - E(Y)^2$$

$$= \lambda^2 + \lambda - \lambda^2 = \lambda$$

3.1 Continous Random Variables

1. Probability Density Function (pdf)

continuous random variable Y

$$\int_a^b f(y)\,dy = P(a \leq Y \leq b)$$

Cumulative probability distribution function (cdf)

$$F(Y) = P(Y \leq y),$$

Properties: (1) $\lim_{y \to -\infty} F(Y) = 0$, $\lim_{y \to \infty} F(Y) = 1$

(2) $F(Y)$ is non-increasing

(3) $F(Y)$ is continous

Relations: $f(y) = F'(y) = \frac{dF(y)}{dy}$

$F(y) = \int_{-\infty}^{y} f(t)\,dt$

Consider: $f(y) = \begin{cases} 3y^2, & 0 \leq y \leq 1 \\ 0, & y \geq 1 \end{cases}$ $f(y)$ is valid

Calculate probabilities: $P(a \leq Y \leq b) = F(b) - F(a) = \int_a^b f(y)\,dy$

3.2 Expectation, Variance, Moment generating function

$E(Y) = \int_{-\infty}^{\infty} y f(y)\, dy$

$E(g(Y)) = \int_{-\infty}^{\infty} g(y) f(y)\, dy$ } both absolute converge

$Var(Y) = E((Y-\mu)^2) = \int_{-\infty}^{\infty} (y-\mu)^2 f(y)\, dy$

$m(t) = E(e^{tY})$

moment generating functions: uniqueness

If $m_X(t) = m_Y(t)$, for $t \in U(0, \delta)$

Then $\forall u$, $F_X(u) = F_Y(u)$

3.3 Chebyshev's Inequality

$P(|Y-\mu| < k\sigma) \geq 1 - \frac{1}{k^2}$

$Var(Y) = \sigma^2 = \int_{-\infty}^{\infty} (y-\mu)^2 f(y)\, dy$

$\geq \int_{-\infty}^{\mu-k\sigma} (y-\mu)^2 f(y)\, dy + \int_{\mu+k\sigma}^{\infty} (y-\mu)^2 f(y)\, dy$

$\geq \int_{-\infty}^{\mu-k\sigma} k^2\sigma^2 f(y)\, dy + \int_{\mu+k\sigma}^{\infty} k^2\sigma^2 f(y)\, dy$

$= k^2\sigma^2 \cdot [P(Y \leq \mu-k\sigma) + P(Y \geq \mu+k\sigma)]$.

3.4 Distributions

(1) Uniform Distribution

$$f(y) = \begin{cases} \frac{1}{b-a}, & a \le y \le b \\ 0, & \text{elsewhere} \end{cases}$$

$$Y \sim U(a,b)$$

$$E(Y) = \frac{a+b}{2}, \quad Var(Y) = \frac{(a-b)^2}{12}, \quad m_Y(t) = \frac{e^{tb} - e^{ta}}{t(b-a)}$$

(2) Normal Distribution

$$f(y) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(y-\mu)^2}{2\sigma^2}}$$

$$Y \sim N(\mu, \sigma^2)$$

Standard normol distribution : $N(0,1)$

$$P(a \le Y \le b) = P\left(\frac{a-\mu}{\sigma} \le Z \le \frac{b-\mu}{\sigma}\right)$$

$$Y \sim N(\mu, \sigma^2), \quad Z \sim N(0,1)$$

$$E(Y) = \mu, \quad Var(Y) = \sigma^2, \quad m_Y(t) = e^{\mu t + \frac{\sigma^2 t^2}{2}}$$

(3) Exponential Distribution

$$f(y) = \begin{cases} \lambda e^{-\lambda y}, & 0 \le \lambda < \infty, \lambda > 0 \\ 0, & \text{otherwise} \end{cases}$$

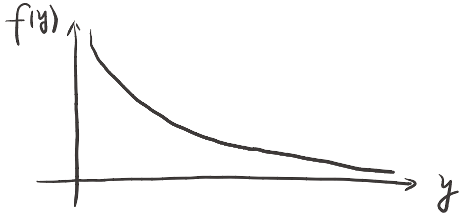
$$E(Y) = \frac{1}{\lambda}, \quad Var(Y) = \left(\frac{1}{\lambda}\right)^2, \quad m_Y(t) = \frac{\lambda}{\lambda - t}$$

Property : $P(Y > a+b \mid Y > a) = P(Y > b)$

Example. $Y \sim \exp(\frac{1}{24})$

$$P(Y \le 12) = \int_0^{12} \frac{1}{24} e^{-\frac{y}{24}} dy = -e^{-\frac{y}{24}}\Big|_0^{12} = 1 - e^{-\frac{1}{2}}$$

$$P(Y > 36) = 1 - P(Y \le 36) = 1 - \int_0^{36} \frac{1}{24} e^{-\frac{y}{24}} dy = e^{-\frac{3}{2}}$$



$$\int_0^m \frac{1}{24}\cdot e^{-\frac{y}{24}}\,dy = 0.5 \Rightarrow e^{-\frac{y}{24}}\Big|_0^m = 0.5 \Rightarrow m = 24\log 2$$

Hazard Function: $h_Y(y) = \lim_{\delta \to 0} \frac{P(y \le Y \le y+\delta \mid Y \ge y)}{\delta} = \frac{f(y)}{1-F(y)}$

(4) Gamma Distribution

$$f(y) = \begin{cases} \frac{\lambda e^{-\lambda y}(\lambda y)^{\alpha-1}}{\Gamma(\alpha)}, & y \geq 0,\ \alpha, \lambda > 0 \\ 0, & y < 0 \end{cases}$$

where $\Gamma(\alpha) = \int_0^\infty y^{\alpha-1} e^{-y}\,dy$

Properties.
- $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$
- $\Gamma(1) = 1$
- If α is an integer, then $\Gamma(n) = (n-1)!$

$$\Gamma(\alpha) = \int_0^\infty y^{\alpha-1} e^{-y}\,dy = -e^{-y}y^{\alpha-1}\Big|_0^\infty + \int_0^\infty e^{-y}(\alpha-1)y^{\alpha-2}\,dy$$
$$= 0 + (\alpha-1) + \int_0^\infty y^{\alpha-2} e^{-y}\,dy = (\alpha-1) + \Gamma(\alpha-1)$$

$E(Y) = \frac{\alpha}{\lambda}$, $Var(Y) = \frac{\alpha}{\lambda^2}$, $m_Y(t) = \frac{1}{(1-\frac{t}{\lambda})^\alpha}$

(5) chi-square Distribution

$$f(y) = \begin{cases} \frac{y^{\frac{\nu}{2}-1} e^{-\frac{y}{2}}}{2^{\frac{\nu}{2}}\Gamma(\frac{\nu}{2})}, & 0 \le y < \infty,\ \nu > 0 \\ 0, & \text{elsewhere} \end{cases}$$

ν: degrees of freedom

3.5 Poisson Process

time between events $i-1$ and i by X_i, $i=1$:

assume that $X_i \overset{iid}{\sim} \exp(\lambda)$

$\{N(t) \geq n\} = \{S_n \leq t\}$

$$P(N(t)=n) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}$$

$$P(S_n \leq t) = \int_0^t \lambda e^{-\lambda y} \frac{(\lambda y)^{n-1}}{(n-1)!} dy$$

3.6 Transformation of random variables

Y: random variable U: a function of Y

$U(Y)$ distribution

1. CDF Method (Culmulative distribution functions)

Example 1. $Y = \begin{cases} 2y, & 0 \le y \le 1 \\ 0, & \text{elsewhere} \end{cases}$ $\qquad U = 3Y - 1$

$$f_U(u) = P(U \le u) = P(3Y - 1 \le u) = P(Y \le \tfrac{u+1}{3})$$

$$= \int_{-\infty}^{\frac{u+1}{3}} f(y)\,dy = \int_0^{\frac{u+1}{3}} 2y\,dy = \frac{(u+1)^2}{9}$$

$$f_U(u) = \begin{cases} 0, & u < -1 \\ \frac{(u+1)^2}{9}, & -1 \le u \le 2 \\ 1, & u > 2 \end{cases}$$

$$\frac{d}{du} f_U(u) = \begin{cases} \frac{2(u+1)}{9}, & -1 \le u \le 2 \\ 0, & \text{elsewhere} \end{cases}$$

Example 2. $Y \sim U(0,1)$, $U = Y^2$, pdf of U?

$f(y) = \begin{cases} 1, & 0 \le y \le 1 \\ 0, & \text{elsewhere} \end{cases}$ $\qquad$ $0 \le y \le 1$, $0 \le y^2 \le 1$, $0 \le u \le 1$

$$P(U \le u) = P(Y^2 \le u) = P(Y \le \sqrt{u}) = \int_{-\infty}^{\sqrt{u}} f(y)\,dy$$

$$= \int_0^{\sqrt{u}} dy = \sqrt{u}$$

$$F_U(u) = \begin{cases} 0, & u < 0 \\ \sqrt{u}, & 0 \le u \le 1 \\ 1, & u > 1 \end{cases}, \quad f_U(u) = \begin{cases} \frac{1}{2\sqrt{u}}, & 0 \le u \le 1 \\ 0, & \text{elsewhere} \end{cases}$$

2. Tranform Method

(1) Increasing / Decreasing Functions

(2) Method

$f_Y(y)$: pdf

$U = h(Y)$ pdf : $f_U(u) = f_Y(h^{-1}(u)) \left| \dfrac{dh^{-1}(u)}{du} \right|$

I regions, each increasing : $f_U(u) = f_Y(h_i^{-1}(u)) \cdot \left| \dfrac{dh_i^{-1}(u)}{du} \right|$

Example 1. $f(y) = \begin{cases} 2y, & 0 \le y \le 1 \\ 0, & \text{elsewhere} \end{cases}$, $U = h(Y) = 4Y+3$

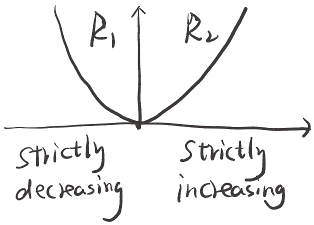
$\boxed{\begin{array}{l} 0 \le y \le 1 \\ 3 \le u \le 7 \end{array}}$

$u = 4y+3$, $y = \dfrac{u-3}{4}$

$h^{-1}(u) = \dfrac{u-3}{4}$ $\left| \dfrac{dh^{-1}(u)}{du} \right| = \left| \dfrac{1}{4} \right| = \dfrac{1}{4}$

$f_U(u) = 2y = 2 \cdot \dfrac{u-3}{4} \cdot \dfrac{1}{4} = \dfrac{u-3}{8}$

Example 2 $Y \sim N(0,1)$, $U = Y^2$, pdf of U?



$R_1 : y < 0$ $u = y^2$, $y = -\sqrt{u}$, $h_1^{-1}(u) = -\sqrt{u}$, $\left| \dfrac{d}{du} h_1^{-1}(u) \right| = \left| -\dfrac{1}{2\sqrt{u}} \right| = \dfrac{1}{2\sqrt{u}}$

$R_2 : y > 0$ $u = y^2$, $y = \sqrt{u}$, $h_2^{-1}(u) = \sqrt{u}$, $\left| \dfrac{d}{du} h_2^{-1}(u) \right| = \dfrac{1}{2\sqrt{u}}$

$f_Y(y) = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}$,

$f_U(u) = \dfrac{1}{\sqrt{2\pi}} \cdot e^{-\frac{(-\sqrt{u})^2}{2}} \dfrac{1}{2\sqrt{u}} + \dfrac{1}{\sqrt{2\pi}} \cdot e^{-\frac{(\sqrt{u})^2}{2}} \cdot \dfrac{1}{2\sqrt{u}}$

$= \dfrac{e^{-\frac{u}{2}} u^{-\frac{1}{2}}}{\sqrt{2}\sqrt{\pi}}$ chi : 1df

Gamma : $\alpha = \frac{1}{2}$

3. MGF Method

$Y_1, Y_2 \quad m_{Y_1,Y_2}(t_1,t_2) = E(e^{t_1Y_1+t_2Y_2})$

(1) Sum of Independent Variables

If $U = \sum_{i=1}^{n} Y_i$ and independent, then $m_U(t) = m_{Y_1}(t)\, m_{Y_2}(t) \cdots m_{Y_n}(t)$

Proof. $m_U(t) = E(e^{t_1Y_1+\cdots+t_nY_n}) = E(e^{t_1Y_1}) \cdots E(e^{t_nY_n})$

(2) Method

$m_U(t)$ and $m_V(t)$ has the same expression, V is already known, then $U=V$.

Example Y_1, Y_2 independent

$Y_1 \sim Poisson(\lambda_1)$, $Y_2 \sim Poisson(\lambda_2)$

$$m_X(t) = m_{Y_1}(t)\, m_{Y_2}(t) = e^{\lambda_1(e^t-1)} e^{\lambda_2(e^t-1)}$$
$$= e^{(\lambda_1+\lambda_2)(e^t-1)}$$

Hence, $X \sim Poisson(\lambda_1+\lambda_2)$

4. Order Statistics

order the random variables :

$Y_1, Y_2, \cdots, Y_n$ with cdf $F(Y)$, pdf $f(y)$

$Y_{(1)}, Y_{(2)}, \cdots, Y_{(n)}$ such that $Y_{(1)} \leq Y_{(2)} \leq \cdots \leq Y_{(n)}$

(1) Distribution of maximum : (cdf method)

$$F_{Y_{(n)}}(y) = P(Y_{(n)} \leq y) = P(Y_1 \leq y \cap Y_2 \leq y \cap \cdots \cap Y_n \leq y)$$
$$= P(Y_1 \leq y)\, P(Y_2 \leq y) \cdots P(Y_n \leq y) \text{ (independent)}$$
$$= F_Y(y) \cdot F_Y(y) \cdots F_Y(y) = [F_Y(y)]^n$$
$$f_{Y_{(n)}}(y) = \frac{d}{dy} F_{Y_{(n)}}(y) = n \cdot [F_Y(y)]^{n-1} f_Y(y)$$

(2) Distribution of minimum : (cdf method)

$$F_{Y_{(1)}}(y) = P(Y_{(1)} \leq y) = 1 - P(Y_{(1)} > y)$$
$$= 1 - P(Y_{(1)} > y)\, P(Y_{(2)} > y) \cdots P(Y_{(n)} > y)$$
$$= 1 - [1 - F_Y(y)]^n$$
$$f_{Y_{(1)}}(y) = \frac{d}{dy} F_{Y_{(1)}} y = n [1 - F_Y(y)]^{n-1} f_Y(y)$$

(3) the kth order

$$f_{Y_{(k)}}(y) = \frac{n!}{(k-1)!(n-k)!} (F_Y(y))^{k-1} (1 - F_Y(y))^{n-k} f_Y(y)$$

3.7 Simulating random variables

Inverse Transform Method

1) Simulate u from $U(0,1)$
2) Find CDF $F_X^{-1}(x)$
3) Set $X = F_X^{-1}(u)$

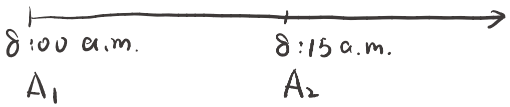
Example $F_X^{-1}(x) = -\frac{\log(1-x)}{\lambda}$, $X = F_X^{-1}(u) = -\frac{\log(1-u)}{\lambda}$

$$f(x) = \lambda e^{-\lambda x}, x>0, \lambda>0$$

$$F(x) = 1 - e^{-\lambda x} \Rightarrow e^{-\lambda x} = 1-y \ , -\lambda x = \log(1-y)$$

Motivate Question Answers

Y : time to interview applicant , $Y \sim exp(2)$



$$P(Y > \tfrac{1}{4}) = 1 - P(Y \le \tfrac{1}{4}) = 1 - \int_0^{\frac{1}{4}} 2e^{-2y}\,dy = 0.6065 \quad //$$

4.1 Discrete Bivariate Distribution

$Y_1 \quad S_1$

$Y_2 \quad S_2$

joint pmf $f(y_1, y_2)$ for Y_1, Y_2

$P(y_1, y_2) = P\{(Y_1 = y_1) \cap (Y_2 = y_2)\}$

Example Tossing a coin 3 times

Y_1: heads for first two times
Y_2: tails for total three times

S	y_1	y_2
HHH	2	0
HHT	2	1
HTH	1	1
THH	1	1
HTT	1	2
THT	1	2
TTH	0	2
TTT	0	3

- Properties: $P(x, y) \geq 0 \qquad \sum_x \sum_y P(x, y) = 1$

Marginal probability mass function:

$$f_1(x) = \sum_y P(x, y)$$

$$f_2(y) = \sum_x P(x, y)$$

- Conditional Probability

$$P(y_1 | y_2) = P(Y_1 = y_1 | Y_2 = y_2) = \frac{P(Y_1 = y_1, Y_2 = y_2)}{P(Y_2 = y_2)}$$

$$= \frac{P(y_1, y_2)}{f_2(y_2)}$$

4.2 Multinomial Distribution

An experiment:

n identical trials. K outcomes, independent

Y_i: the number that result in outcome i

$\sum_{i=1}^{k} Y_i = n$,

PMF: $P(y_1, \cdots, y_k) = \frac{n!}{y_1! \cdots y_k!} p_1^{y_1} \cdots p_k^{y_k}$, $\sum_{i=1}^{k} p_k = 1$

Properties:

$E(Y_i) = np_i$

$Var(Y_i) = np_i(1-p_i)$

$cov(Y_i, Y_j) = -np_i p_j$, $i \neq j$

Covariance: $cov(Y_1, Y_2) = E[(Y_1 - \mu_1)(Y_2 - \mu_2)]$, $\mu_i = E(Y_i)$

4.3 Continuous Bivariate Distributions

PDF: $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dx\,dy = 1$, $f(x,y) \geq 0$, $\forall -\infty < x, y < \infty$

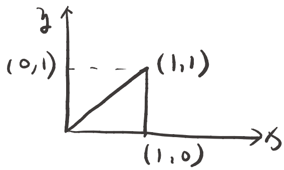
$P(a_1 \leq X \leq a_2,\ b_1 \leq Y \leq b_2) = \int_{a_1}^{a_2}\int_{b_1}^{b_2} f(x,y)\,dy\,dx$

Example 1 $f(x,y) = \begin{cases} 1, & 0 \leq x \leq 1,\ 0 \leq y \leq 1 \\ 0, & \text{else} \end{cases}$

Region Integration: $\int_0^1\int_0^1 dx\,dy = \int_0^1 dy = 1$

$P(0.1 \leq x \leq 0.3,\ 0 \leq y \leq 0.5) = 0.1$

Example 2. $f(x,y) = \begin{cases} 2, & 0 \leq y \leq x \leq 1 \\ 0, & \text{else} \end{cases}$



$\int_0^1\int_y^1 2\,dx\,dy = 1$, so it is valid

Marginal Probability function

$f_1(x) = \int_{-\infty}^{\infty} f(x,y)\,dy$

$f_2(y) = \int_{-\infty}^{\infty} f(x,y)\,dx$

Example 3. $f(x,y) = \begin{cases} 2x, & 0 \leq x \leq 1,\ 0 \leq y \leq 1 \\ 0, & \text{else} \end{cases}$

marginal of X and Y:

$f_1(x) = \int_0^1 2x\,dy = 2xy\Big|_0^1 = 2x,\ 0 \leq x \leq 1$

$f_2(y) = \int_0^1 2x\,dx = x^2\Big|_0^1 = 1,\ 0 \leq y \leq 1$

4.4 Conditional PDF

$$f(x|y) = \frac{f(x,y)}{f_2(y)} \longrightarrow$$

marginal PDF

$$f(y|x) = \frac{f(x,y)}{f_1(x)} \longrightarrow$$

Example 1

$$f(x,y) = \begin{cases} 2x, & 0 \le x \le 1,\ 0 \le y \le 1 \\ 0, & \text{else} \end{cases}$$

$$f(x|y) = \frac{f(x,y)}{f_2(y)} = \frac{2x}{1} = 2x, \quad 0 \le x \le 1$$

$$f(y|x) = \frac{f(x,y)}{f_1(x)} = \frac{2x}{2x} = 1, \quad 0 \le y \le 1$$

Independence

X and Y are independent $\Leftrightarrow F(x,y) = F_1(x)\,F_2(y)$

(both for discrete and continuous variables)

Example 2. $f(x,y) = \begin{cases} 6xy^2, & 0 \le x \le 1,\ 0 \le y \le 1 \\ 0, & \text{else} \end{cases}$

$$f_1(x) = \int_0^1 6xy^2\,dy = 2xy^3\Big|_0^1 = 2x$$

$$f_2(y) = \int_0^1 6xy^2\,dx = 3x^2y^2\Big|_0^1 = 3y^2$$

$f_1(x) \cdot f_2(y) = 6xy^2 = f(x,y)$, independent

Example 3. $f(x,y) = \begin{cases} 2, & 0 \le y \le x \le 1 \\ 0, & \text{else} \end{cases}$

$$f_1(x) = \int_0^x 2\,dy = 2x$$

$$f_2(y) = \int_y^1 2\,dx = 2-2y$$

$f_1(x) \cdot f_2(y) = 4x(1-y) \ne f(x,y)$, dependent

4.5 Expectations

$$E[g(X,Y)] = \sum_X \sum_Y g(x,y)\, p(x,y) \quad , \text{ discrete}$$

$$E[g(X,Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y)\, f(x,y)\, dx\, dy \quad , \text{ continuous}$$

X and Y are independent:

$$E[g(X)h(Y)] = E[g(X)] \cdot E[h(Y)]$$

- Covariance

$$\text{cov}[X,Y] = E[(X-\mu_1)(Y-\mu_2)]$$

Properties of Covariance:

$$\text{cov}(X,Y) = E(XY) - E(X)E(Y)$$

$$\text{cov}(X,Y) = \text{cov}(Y,X)$$

$$\text{cov}(X,X) = \text{Var}(X)$$

$$\text{cov}(X,Y) = 0 \underset{\nRightarrow}{\Longleftarrow} X, Y \text{ are independent}$$

- Correlation

$$\rho = \text{corr}(X,Y) = \frac{\text{cov}(X,Y)}{\text{sd}(X)\,\text{sd}(Y)}$$

Linear Combination:

$$E(aX+bY) = aE(X) + bE(Y)$$

$$\text{Var}(aX+bY) = a^2\,\text{Var}(X) + b^2\,\text{Var}(Y) + 2ab\,\text{cov}(X,Y)$$

- MGF

$$m_{(X,Y)}(t_1,t_2) = E\left(e^{t_1 X + t_2 Y}\right)$$

$$m_U(t) = m_{Y_1}(t) \cdots m_{Y_n}(t) \quad (Y_1, \cdots, Y_n \text{ are independent})$$

4.6 Bivariate Normal Distribution

$E(X) = \mu_X, \quad E(Y) = \mu_Y$

$Var(X) = \sigma_X^2, \quad Var(Y) = \sigma_Y^2$

$cov(X, Y) = \rho \sigma_X \sigma_Y$

PDF:

$$f(x, y) = \frac{1}{2\pi \sigma_X \sigma_Y \sqrt{1-\rho^2}} e^{-\frac{Q}{2}}$$

$$Q = \frac{1}{1-\rho^2}\left[\frac{(x-\mu_X)^2}{\sigma_X^2} + \frac{(y-\mu_Y)^2}{\sigma_Y^2} - 2\rho \frac{(x-\mu_X)(y-\mu_Y)}{\sigma_X \sigma_Y}\right]$$

Denote:

$$\begin{bmatrix} X \\ Y \end{bmatrix} \sim N_2(\mu, \Sigma)$$

$$\mu = \begin{bmatrix} \mu_X \\ \mu_Y \end{bmatrix}, \quad \Sigma = \begin{bmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{bmatrix}$$

Marginal Distribution:

$X \sim N(\mu_X, \sigma_X^2), \quad Y \sim N(\mu_Y, \sigma_Y^2)$

Conditional Distribution:

$$Y \mid X = x \sim N\left(\mu_Y + \frac{\rho\sigma_Y}{\sigma_X}(x - \mu_X), \ \sigma_Y^2(1-\rho^2)\right)$$

Multivariate Normal PDF:

$$\text{PDF}: \frac{1}{(2\pi)^{\frac{k}{2}} |\Sigma|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(\vec{x} - \vec{\mu})^T \Sigma^{-1} (\vec{x} - \vec{\mu})\right)$$

$$\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \vec{\mu} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} \quad [\Sigma]_{ij} = cov(X_i, X_j)$$

Linear Combination:

$$U_1 = \sum_{i=1}^{n} a_i Y_i, \quad U_2 = \sum_{j=1}^{m} b_j X_j$$

$$E(U_1) = \sum_{i=1}^{n} a_i \mu_i$$

$$Var(U_1) = \sum_{i=1}^{n} a_i^2 Var(Y_i) + \sum_{i=1}^{n} \sum_{\substack{j=1 \\ i \neq j}}^{n} a_i a_j cov(X_i, X_j)$$

$$cov(U_1, U_2) = \sum_{i=1}^{n} \sum_{j=1}^{m} a_i b_j cov(Y_i, X_j)$$

- Weak Law:

If $X_1, \cdots, X_j$ independent, $E(X_i) = \mu$, $Var(X_i) = \sigma^2$

$$\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$$

$$\forall \varepsilon > 0, \lim_{n \to \infty} P(|\overline{X}_n - \mu| > \varepsilon) = 0$$

- Central Limit Theorem

U_n converges to the standard normal distribution ($n \to \infty$)

$$\forall u, \lim_{n \to \infty} P(U_n \leq u) = \int_{-\infty}^{u} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

4.7 Conditional Expectation

$$E[g(X) \mid Y=y] = \sum_x g(x) P(x \mid y)$$

$$E[g(X) \mid Y=y] = \int_{-\infty}^{\infty} g(x) p(x \mid y)\, dx$$

Example.

$$f(x,y) = \begin{cases} \frac{1}{2}, 0 \le x \le y \le 2 \\ 0, \text{ else} \end{cases}$$

$$f_2(y) = \int_0^y \frac{1}{2}\, dx = \frac{1}{2} y, \quad 0 \le y \le 2$$

$$f(x \mid y) = \frac{f(x,y)}{f_2(y)} = \frac{1}{y}, \quad 0 \le x \le y \le 2$$

$$E[X \mid Y=y] = \int_0^y x \cdot \frac{1}{y}\, dx = \frac{1}{2} y$$

$$E[X \mid Y=1.5] = \frac{1.5}{2} = \frac{3}{4}$$

· Conditional Expectation Theorem

$$E[X] = E[E[X \mid Y]]$$

$$\begin{aligned} E[X] &= \sum_x x P(X=x) = \sum_x x \sum_y P(X=x, Y=y) \\ &= \sum_x x \sum_y P(X=x \mid Y=y) \cdot P(Y=y) \\ &= \sum_x \sum_y x P(X=x \mid Y=y) P(Y=y) \\ &= \sum_y \sum_x x P(X=x \mid Y=y) P(Y=y) \\ &= \sum_y E[X \mid Y=y] P(Y=y) \\ &= E_Y[E[X \mid Y]] P(Y=y) \end{aligned}$$

For variance:

$$Var(X) = E[Var(X \mid Y)] + Var(E[X \mid Y])$$

Motivation Example.

Y : eggs laid $\quad Y \sim Po(15)$

X : eggs hatch $\quad X \mid Y \sim Bin(Y, 0.001)$

$$E[X] = E[E[X \mid Y]] = E[0.001 Y] = 0.001 E[Y]$$
$$= 0.015$$

$$Var(X) = E[Var(X \mid Y)] + Var(E[X \mid Y])$$
$$= E[0.001 \cdot 0.999 \cdot Y] + Var[0.001 Y]$$
$$= 0.001 \cdot 0.999 \cdot E[Y] + 0.001^2 \cdot Var(Y)$$
$$= 0.001 \cdot 0.999 \cdot 15 + 0.001^2 \cdot 15$$
$$= 0.015$$

3.1 Markov Chain Notations

Random Process : $\{X_n\}_{n \in T}$ Discrete : $T = \{0, 1, 2, \cdots, \}$

State Space : Countable

Discrete-time Markov Chain :

$$P(X_n = s \mid X_0 = \gamma_0, \cdots, X_{n-1} = \gamma_{n-1}) = P(X_n = s \mid X_{n-1} = \gamma_{n-1})$$

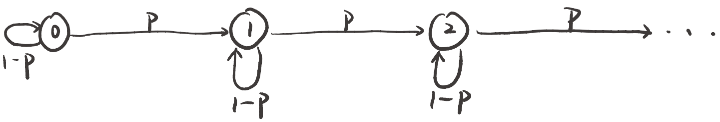
for all $n \geq 1$, all $s, \gamma_0, \cdots, \gamma_{n-1} \in S$

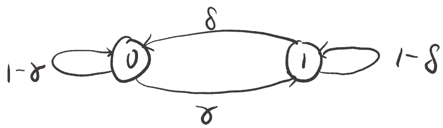
Time-homogeneous Markov Chain

$$P(X_{n+1} = j \mid X_n = i) = P(X_1 = j \mid X_0 = i)$$

$$p_{ij} = P(X_{n+1} = j \mid X_n = i)$$

State Transition Diagram



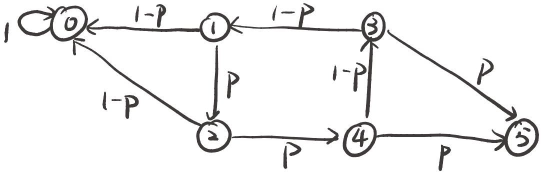


5.2 Transition Matrices

Transition Matrix :

$$[\mathbb{P}]_{ij} = P_{ij} = P(X_{n+1}=j \mid X_n = i)$$

Example



$$\begin{array}{c|c} & \begin{array}{cccccc} 0 & 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} & \begin{pmatrix} 1 & & & & & \\ 1-p & & & & & \\ & & & & & \\ & & & & & \\ & & & & & \\ & & & & & \end{pmatrix} \end{array}$$

m-step transition matrix :

$$[\mathbb{P}]_{ij} = p_{ij}^{(m)} = P(X_{n+m}=j \mid X_n = i)$$

Formula: $\mathbb{P}^{(m)} = \mathbb{P}^m$, for $m \in \mathbb{N}^+$

Example. $\mathbb{P} = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}$

$$[\mathbb{P}^{(3)}]_{(1,2)} = \mathbb{P}^3(1,2)$$

Vector $\vec{p}(0) = (P(X_0=0), P(X_0=1), \cdots)$

$\vec{p}(n) = \vec{p}(0)\,\mathbb{P}^n$ (Distribution of states)

5.3 Equilibrium Distribution

Assume that : $\lim_{m\to\infty} P_{ij}^{(m)} = \pi_j$

Equilibrium Equations : $\vec{\pi} = \vec{\pi}\cdot\mathbb{P}$, $\vec{\pi} = (\pi_0, \pi_1, \pi_2, \cdots)$

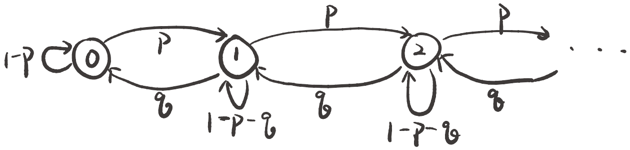
$$P_{ij}^{(m)} = \sum_{k\in S} P_{ik}^{(m-1)} P_{kj}$$

$$\Rightarrow \lim_{m\to\infty} P_{ij}^{(m)} = \lim_{m\to\infty} \sum_{k\in S} P_{ik}^{(m-1)} P_{kj}$$

$$\Rightarrow \pi_j = \sum_{k\in S} \pi_k \cdot P_{kj}$$

Property : $\sum_{i\in S} \pi_i = 1$

Example : Queue



$$\pi_0 = \pi_0(1-p) + \pi_1 q$$

$$\pi_i = \pi_{i-1} p + \pi_i(1-p-q) + \pi_{i+1} q$$

Stationary : $\vec{\pi}\cdot\mathbb{P} = \vec{\pi}$,

$\vec{\pi}$ is the equilibrium probability distribution (ergodic

5.4 Solving Equilibrium Distributions

Finite number of cases:

N linear equation, N unknowns

Finite Queue: $S = \{0, 1, 2, \cdots, N\}$

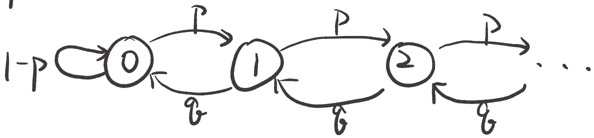
$$p_{i,i+1} = p, \quad 0 \le i < N$$
$$p_{i,i-1} = q, \quad 0 \le i < N$$
$$p_{i,i} = 1-p-q \quad 0 < i < N$$
$$p_{0,0} = 1-p$$
$$p_{N,N} = 1-q$$

Equations:
$$\begin{cases} \pi_0 = \pi_1 q + \pi_0(1-p) \\ \pi_i = \pi_{i+1} q + \pi_i(1-p-q) + \pi_{i-1} p, \quad 0 < i < N \\ \pi_N = \pi_{N-1} p + \pi_N (1-q) \end{cases}$$

$$\pi_0 = \pi_1 q + \pi_0(1-p)$$
$$\pi_1 = \frac{p}{q}\pi_0$$
$$\pi_1 = \pi_2 q + \pi_1(1-p-q) + \pi_0 \cdot q$$
$$q\pi_2 = (p+q)\cdot\frac{p}{q}\pi_0 - p\pi_0, \quad \pi_2 = \left(\frac{p}{q}\right)^2 \pi_0$$
$$\cdots, \quad \pi_i = \left(\frac{p}{q}\right)^i \pi_0$$
$$\sum_{i=0}^{N} \pi_i = 1, \quad \sum_{i=0}^{N} \left(\frac{p}{q}\right)^i \pi_0 = 1, \quad \pi_0 \cdot \frac{1-\left(\frac{p}{q}\right)^{N+1}}{1-\frac{p}{q}} = 1$$
$$\Rightarrow \pi_0 = \frac{1-\frac{p}{q}}{1-\left(\frac{p}{q}\right)^{N+1}}$$
$$\pi_i = \frac{1-\frac{p}{q}}{1-\left(\frac{p}{q}\right)^{N+1}} \cdot \left(\frac{p}{q}\right)^i$$

Infite States :

Infite Queue



$$x_i = p x_{i-1} + (1-p-q) x_i + q x_{i+1} \ , \ i \geq 1$$

$$x_0 = (1-p) x_0 + q x_1$$

$$(p+q) x_i = p x_{i-1} + q x_{i+1}$$

Assume $x_i = m^i$, $(p+q) m^i = p m^{i-1} + q m^{i+1}$

$$\Rightarrow q m^2 - (p+q) m + p = 0$$

$$\Rightarrow (m-1)(qm - p) = 0$$

$$\Rightarrow m = 1 \text{ or } m = \frac{p}{q}$$

$$x_i = a \cdot \left(\frac{p}{q}\right)^i + b \cdot 1^i = a \cdot \left(\frac{p}{q}\right)^i + b$$

$$x_0 = (1-p) x_0 + q x_1 \Rightarrow p x_0 = q x_1$$

$$p(a+b) = q\left(\frac{ap}{q} + b\right) \Rightarrow ap + bp = ap + bq \Rightarrow b(p-q) = 0 \ , b=0$$

$x_i = a \cdot \left(\frac{p}{q}\right)^i$:

$$\sum_{i=0}^{\infty} x_i = 1 \Rightarrow \sum_{i=0}^{\infty} a \cdot \left(\frac{p}{q}\right)^i = 1 \ . \ a \cdot \frac{1}{1 - \frac{p}{q}} = 1 \ , \ a = 1 - \frac{p}{q}$$

$$\Rightarrow x_i = \left(1 - \frac{p}{q}\right) \cdot \left(\frac{p}{q}\right)^i$$