Probability Theory

content

1 Basical Probabilistic Model

1.1 Space of Elementary Events

if a test has limited consequences , these consequences $\omega_1, \cdots, \omega_N$ are called **Elementary Events** , and

$$\Omega = \{\omega_1, \cdots, \omega_N\}$$

is called **(Limited) Space of Elementary Events**

if A$\subseteq\Omega$, A is an **Event**
if A$\subseteq$ Ω and B$\subseteq$ Ω , then

1.2 Classcal Probabilistic Models

suppose $\omega_1, \cdots, \omega_N \subseteq \Omega$ and $N < \infty$, then

$$P(\omega_i) = \frac{1}{N}$$

$\forall A \in \mathscr{A}$

$$P(A) = \frac{N(A)}{N}$$

1.2.1 Random Sampling

Order Sampling with Replacement choose N balls from M boxes is a way of order sampling with replacement

mark the consequence event as A , and $A = (a_1, \cdots, a_N)$

$$\Omega = \{\omega : \omega = (a_1, \cdots, a_n), a_i = 1, \cdots, M\}$$

and

$$N(\Omega) = M^n$$

Disorder Sampling with Replacement if $N < M$

$$\Omega = \{\omega : \omega = [a_1, \cdots, a_n], a_i = 1, \cdots, M\}$$

and

$$N(\Omega) = C_{M+N-1}^{N}$$

Order Sampling without Replacement

$$\Omega = \{\omega : \omega = (a_1, \cdots, a_N), a_k \neq a_l, k \neq l, a_i = 1, \cdots, M\}$$

and

$$N(\Omega) = A_M^N$$

Disorder Sampling without Replacement

$$\Omega = \{\omega : \omega = [\omega_1, \cdots, \omega_N], a_k \neq a_l, k \neq l, a_i = 1, \cdots, N\}$$

and

$$N(\Omega) = C_M^N$$

1.2.2 Arrangment

1.2.3 Binomial Distribution

toss a coin n times in a row , and for $(a_1, \cdots, a_N)$, when it is obverse side $a_i = 1$ and it is reverse side $a_i = 0$

$$\Omega = \{\omega : \omega = (a_1, \cdots, \omega_n), a_i = 0 or 1\}$$

and

$$\forall a_i, P(a_i = 1) = p$$

then

$$P(\omega) = p^{\sum a_i}(1-p)^{n-\sum a_i}$$

$$P(A) = \sum_{\omega \in A} P(\omega) \quad A \in \mathscr{A}$$

1.2.4 Multinomial Distribution

1.2.5 Hypergeoetric Distrbution

e.g there are M balls in a box , their number are $1, \cdots, M$, and M_i balls have color c_i , $\sum M_i = M$, choose n balls ,the number of balls that color is c_i is n_i ,then

$$P(B_{n_1,\cdots,n_r}) = \frac{C_{M_1}^{n_1} \cdots C_{M_r}^{n_r}}{C_M^n}$$

1.2.6 Stirling's Approximation

$$n! = \sqrt{2\pi n}(\frac{n}{e})^n e^{\frac{\theta_n}{12n}} \quad ,0 < \theta_n < 1$$

1.3 Conditional Probability

1.3.1 Conditional Probbility

the probability of B under the condition of A is **Conditional Probability** , marked as

$$P(B \mid A)$$

and

$$P(B \mid A) = \frac{P(AB)}{P(A)}$$

1.3.2 Total Probability Theorem

suppose $A_1 + \cdots + A_n = \Omega$, $B \in \Omega$

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

1.3.3 Bayes Rule

if $A, B \in \Omega$, $P(A), P(B) > 0$

$$P(A_i) = \frac{P(A_i)P(B \mid A_i)}{\sum_{j=1}^{n} P(A_j)P(B \mid A_j)}$$

1.3.4 Independence

if A and B are independent

$$P(AB) = P(A)P(B)$$

2 Random Variable

2.1 Cumulative Distribution Function

$$\forall x, F(x) = P\{X \leq x\} = P\{\omega \mid X(\omega) \leq x\}$$

Therom:

$$\forall x_1, x_2, if \quad x_1 \leq x_2, \quad F(x_1) \leq F(x_2)$$

$$0 \leq F(x) \leq 1, \lim_{x \to -\infty} F(x) = 0, \lim_{x \to +\infty} F(x) = 1$$
$$F(x+0) = F(x)$$

2.1.1 Discrete Random Variable

if randdom variable X is limited and the value are $x_1, \cdots, x_n, \cdots$, $P\{X = x_i\} = p_i$, then

$$p_i \geq 0$$
$$\sum_{i=1}^{\infty} p_i = 1$$

then X is a discrete random variable

2.2 Common Distribution

2.2.1 Poission Distribution

$$P\{X = k\} = \frac{\lambda^k}{k!} e^{-\lambda}, k = 0, 1, \cdots; \quad \lambda > 0$$

Poission Therom:

2.3 Probability Density Function

$$F(X) = \int_{-\infty}^{x} f(u)du$$

$f(x)$ is probability density function

2.3.1 Uniform Distribution

$$f(x) = \begin{cases} \frac{1}{b-a}, a < x < b \\ 0\ ,\ other \end{cases}$$

$X \sim U(a,b)$

$$F(x) = \begin{cases} 0\ ,\ x < a \\ \frac{x-a}{b-a}\ ,\ a \leq x < b \\ 1\ ,\ b \leq x \end{cases}$$

2.3.2 Exponential Distribution

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

2.3.3 Normal Distribution

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

$$\phi(x;\mu,\sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\ ,\ x \in R$$

$$\forall a, b\ ,\ X \sim N(a,b) \Rightarrow P(x_1 < x < x_2) = \Phi(\frac{x_2 - a}{\sqrt{b}}) - \Phi(\frac{x_1 - a}{\sqrt{b}})$$

$$\Phi(-x) = 1 - \Phi(x)$$

2.4 Joint Distribution Function

if (X, Y) is in Ω , their distribution function are

$$F_X(x) = P\{X \leq x\} \ , \ F_Y(y) = P\{Y \leq y\}$$

joint distribution function is

$$F(x, y) = P\{X \leq x, Y \leq y\}$$

Therom:

$$P\{X \leq x\} = P\{X \leq x, Y < +\infty\}$$
$$P\{Y \leq y\} = P\{X < +\infty, Y \leq y\}$$
$$F_X(x) = \lim_{y \to +\infty} F(x, y)$$
$$F_Y(y) = \lim_{x \to +\infty} F(x, y)$$
$$P\{x_1 < X \leq x_2, y_1 < Y \leq y_2\} = F(x_2, y_2) - F(x_2, y_1) - F(x_1, y_2) + F(x_1, y_1)$$

2.4.1 Discrete Jiont Distribution Function

$$F(x, y) = \sum_{x_i \leq x} \sum_{y_j \leq j} p_{ij}$$

2.4.2 Joint Probability Density

$$F(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(u, v) du dv$$
$$G \subset \mathbb{R}^2 \ , \ P\{(X, Y) \in G\} = \iint_G f(x, y) d\sigma$$
$$f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy \ , \ x \in \mathbb{R}$$
$$f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx \ , \ y \in \mathbb{R}$$
$$F_X(x) = F(x, +\infty) = \int_{-\infty}^{x} [\int_{-\infty}^{+\infty} f(u, v) dv] dv$$

2.4.3 Geometric Probability

$$f(x,y)=\begin{cases}\dfrac{1}{S(G)}\ ,\ (x,y)\in G\\ 0\ ,\ (x,y)\notin G\end{cases}$$

2.4.4 Two-Dimensional Normal Distribution

$$\phi(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}e^{-\frac{1}{2(1-\rho^2)}[\frac{(x-\mu_1)^2}{\sigma_1^2}-2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}]}\ ,\ x\in\mathbb{R}\ ,\ y\in\mathbb{R}$$

2.4.5 Independence

$$F(x,y)=F_X(x)F_Y(y)$$
$$f(x,y)=f_X(x)f_Y(y)$$

2.4.6 Conditional Distribution

$$P\{Y=y_j,X=x_i\}=\frac{p_{ij}}{p_i}$$
$$f_{X|Y}(x\mid y)=\frac{f(x,y)}{f_X(x)}$$

2.4.7 Random Variable Function

if $y=y(x)$

$$f_Y(y)=\begin{cases}f_X[x(y)]\mid x(y)'\mid\ ,\ \alpha<y\beta\\ 0\ ,\ other\end{cases}$$

$Z_1=max\{X,Y\}\ ,\ Z_2=min\{X,Y\}$

$$F_{Z_1}=P\{max\{X,Y\le z\}=P\{X\le z,Y\le z\}=P\{X\le z\}P\{Y\le z\}=F_X(z)F_Y(z)\}$$

$$F_{Z_2}=P\{min\{X,Y\}\le z\}=1-P\{min\{X,Y\}>z\}=1-P\{X>z,Y>z\}=1-P\{X>z\P\{Y>z\}$$

3 Numercial Characters

3.1 Expection

3.1.1 Discrete Distribution

$$P\{X = x_i\} = p_i \ , \ E(X) = \sum_{\infty}^{i=1} x_i p_i$$

3.1.2 Continuous Distribution

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

3.1.3 Property

$$\begin{aligned}
&E(C) = C \\
&E(CX) = CE(X) \\
&E(X + Y) = E(X) + E(Y) \\
&E(XY) = E(X)E(Y)
\end{aligned}$$

3.1.4 Random Variable Function

3.2 Variance

$$D(X) = E[X - E(X)]^2$$

3.2.1 Discrete Distribution

$$D(X) = \sum_{i=1}^{\infty} [x_i - E(X)]^2 P\{X = x_i\}$$

3.2.2 Continuous Distribution

$$D(X) = \int_{-\infty}^{+\infty} [x - E(X)]^2 f_X(s) dx$$

3.2.3 Property

$$\begin{aligned}
&D(C) = 0 \\
&D(CX) = C^2 D(X) \\
&D(X \pm Y) = D(X) + D(Y) + \pm 2E\{[X - E(X)][Y - E(Y)]\}
\end{aligned}$$

3.3 Expection and Variance of Common Distribution Function

3.3.1 Binomial Distribution

$$X \sim B(n,p) \Rightarrow E(X) = np \ , \ D(X) = np(1-p)$$

3.3.2 Poisson Distribution

$$X \sim P(\lambda) \Rightarrow \ E(X) = \lambda \ , \ D(X) = \lambda$$

3.3.3 Geometric Distribution

$$X \sim U(a,b) \Rightarrow E(X) = \frac{a+b}{2} \ , \ D(X) = \frac{(b-a)^2}{12}$$

3.3.4 Exponential Distribution

$$E(X) = \frac{1}{\lambda} \ , \ D(X) = \frac{1}{\lambda^2}$$

3.3.5 Normal Distribution

$$X \sim N(\mu, \sigma^2) \Rightarrow E(x) = \mu \ , \ D(x) = \sigma^2$$

3.4 Covariance

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

is the covariance of X and Y

especially

$$D(X) = cov(X, X)$$

therefore

$$D(X \pm Y) = D(X) + D(Y) \pm 2cov(X, Y)$$

Therom:

$$cov(X, Y) = cov(Y, X)$$

$$cov(aX, bY) = abcov(Y, X)$$

$$cov(X_1 + X_2, Y) = cov(X_1, Y) + cov(X_2, Y)$$

$$cov(X, Y) = E(XY) - E(X)E(Y)$$

3.5 Correlation Coefficient

for (X, Y) , $D(X) > 0, D(Y) > 0$

$$\rho_{XY} = \frac{cov(X, Y)}{\sqrt{D(X)D(Y)}}$$

$$\rho_{XY} = E[\frac{X - E(X)}{\sqrt{D(X)}} \frac{Y - E(Y)}{\sqrt{D(Y)}}] = E(X^*Y^*) = cov(X^*, Y^*)$$

Therom:

$$\mid \rho_{XY} \mid \leq 1$$

$$\mid \rho \mid = 1 \ \Leftrightarrow \ \exists b, a \neq 0, P\{Y = aX + b\} = 1$$

3.6 Law of large numbers and Central limit Theorem

3.6.1 Law of large numbers

Chebyshev's Theorem

for random variable X , $E(X)$ and $D(X)$ exist , $\forall \epsilon > 0$

$$P\{\mid X - E(X) \mid \geq \epsilon\} \leq \frac{D(X)}{\epsilon^2}$$

or

$$P\{\mid X - E(X) \mid < \epsilon\} \geq 1 - \frac{D(X)}{\epsilon^2}$$

Law of Large Numbers for sequence of random variables $X_1, \cdots, X_n, \cdots$, $E(X_i)$ exists , $i = 0, 1 \cdots$, $\forall \epsilon < 0$

$$\lim_{x \to \infty} P\{\mid \frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} E(X_i) \mid < \epsilon\} = 1$$

3.6.2 Chebyshev Law of Large num-herd

if sequence of random variable $X_1, \cdots, X_n, \cdots$ are independent , $E(X_i)$ and $D(X_i)$ are exsit , and $D(X_i) < C$, $i = 0, 1, \cdots$, $\forall \epsilon > 0$

$$\lim_{n \to n}\{\mid \frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} E(X_i) \mid < \epsilon\} = 1$$

3.6.3 Wiener-Khinchin Law of Large Numbers

for sequence $X_1, \cdots, X_n, \cdots$ are independent ,and $E(X_i) = \mu$, $i = 0, 1 \cdots$, $\forall \epsilon > 0$

$$\lim_{n \to \infty} P\{\mid \frac{1}{n}\sum_{i=1}^{n} X_i - \mu \mid\} < \epsilon\} = 1$$

3.6.4 Bernoulli Law of Large Num-hers

$$\lim_{n \to n} P\{\mid \frac{m}{n} - p \mid < \epsilon\} = 1$$

3.7 Central Limit Theorom

if sequence $X_1, \cdots X_n, \cdots$ are independent

$$\lim_{n\to\infty} P\{\frac{\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} E(X_i)}{\sqrt{\sum_{i=1}^{x} D(X_i)}}\} = \Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt \ , \ -\infty < x < +\infty$$

if $E(X_i) = \mu$, $D(X_i) = \sigma^2 > 0$, $i = 0, 1, \cdots$

$$\lim_{n\to\infty} P\{\frac{\sum_{i=1}^{x} X_i - n\mu}{\sqrt{n}\sigma} \leq x\} = \Phi(x)$$

3.7.1 De Moivre-Laplace

if $Y_n \sim B(n, p)$, $n = 0, 1, \cdots$, $\forall x$

$$\lim_{n\to\infty} P\{\frac{Y_n - np}{\sqrt{np(1-p)}} \leq x\} = \Phi(x)$$