Basic Concepts of Probability

Basic Properties : Events

1. $(A \cap B)' = A' \cup B'$
2. $(A \cup B)' = A' \cap B'$
3. $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
4. $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
5. $A \cup B = A \cup (B \cap A')$
6. $A = (A \cap B) \cup (A \cap B')$

De Morgan's Law

1. $(A_1 \cup A_2 \cup \dots \cup A_n)' = A_1' \cap A_2' \cap \dots \cap A_n'$
2. $(A_1 \cap A_2 \cap \dots \cap A_n)' = A_1' \cup A_2' \cup \dots \cup A_n'$

Permutation and Combination

$${}_nP_r = n(n-1)(n-2)\dots(n-(r-1)) = \frac{n!}{(n-r)!}$$

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

$${}_nC_r\, r! = {}_nP_r$$

Binomial Coefficient $\rightarrow \binom{n}{r}$

1. $\binom{n}{r} = \binom{n}{n-r}$ for $r = 0, 1, \dots, n$
2. $\binom{n}{r} = \binom{n-1}{r} + \binom{n-1}{r-1}$ for $1 \le r \le n$
3. $\binom{n}{r} = 0$ for $r < 0$ or $r > n$

Basic Properties : Probability

1. $Pr(A) = Pr(A \cap B) + Pr(A \cap B')$
2. $Pr(A_1 \cup A_2 \cup \dots \cup A_n) = \sum_{i=1}^{n} Pr(A_i) - \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} Pr(A_i \cap A_j) + \dots + (-1)^{n+1} Pr(A_1 \cap A_2 \cap \dots \cap A_n)$

Inclusion-Exclusion Principle

Conditional Probability

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

1. $Pr(S|A) = 1$
2. $B_1, B_2, \dots$ are mutually exclusive (disjoint) events

$$\longrightarrow Pr\left(\bigcup_{i=1}^{\infty} B_i \,\middle|\, A\right) = \sum_{i=1}^{\infty} Pr(B_i|A)$$

Multiplication Rule of Probability

$$Pr(A \cap B) = Pr(A)\,Pr(B|A)$$
$$= Pr(B)\,Pr(A|B),$$ providing $Pr(A) > 0$, $Pr(B) > 0$

In general, $Pr(A_1 \cap A_2 \cap \dots \cap A_n)$

$$= Pr(A_1)\,Pr(A_2|A_1)\,Pr(A_3|A_1 \cap A_2) \dots Pr(A_n|A_1 \cap A_2 \cap \dots \cap A_n),$$

providing $Pr(A_1 \cap A_2 \cap \dots \cap A_{n-1}) > 0$.

The Law of Total Probability

Let $A_1, A_2, \dots, A_n$ be a partition of S. Then for any event B,

$$Pr(B) = \sum_{i=1}^{n} Pr(B \cap A_i) = \sum_{i=1}^{n} Pr(A_i)\,Pr(B|A_i)$$

Bayes' Theorem

Let $A_1, A_2, \dots, A_n$ be a partition of S. Then

$$Pr(A_k|B) = \frac{Pr(A_k)\,Pr(B|A_k)}{\sum_{i=1}^{n} Pr(A_i)\,Pr(B|A_i)}$$

$$= \frac{Pr(A_k)\,Pr(B|A_k)}{Pr(B)}$$

Random Variables

Discrete Random Variables		Continuous Random Variables

$f_X(x)$

Discrete Random Variables: probability (mass) function

- Must satisfy:

① $f(x_i) \geq 0 \quad \forall\ x_i$

② $\sum_{i=1}^{\infty} f(x_i) = 1$

Continuous Random Variables: probability density function (p.d.f.)

- Must satisfy:

① $f(x) \geq 0 \quad \forall\ x \in R_X$

② $\int_{R_X} f(x)\, dx = 1$ or $\int_{-\infty}^{\infty} f(x)\, dx = 1$

③ $\forall\ c, d$ s.t. $c < d$, $(c,d) \subset R_X$, $\Pr(c \leq X \leq d) = \int_c^d f(x)\, dx$

Cumulative Distribution Function (CDF) $F_X(x)$

Discrete

$$F(x) = \sum_{t \leq x} f(t) = \sum_{t \leq x} \Pr(X = t)$$

① $\Pr(a \leq X \leq b) = \Pr(X \leq b) - \Pr(X < a)$
$= F(b) - F(a-1)$

② $a = b \longrightarrow \Pr(X = a) = F(a) - F(a-1)$

Continuous

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

① $\Pr(a \leq X \leq b) = \Pr(a < X \leq b)$
$= F(b) - F(a)$

② $f(x) = \frac{d\,F(x)}{dx} \longrightarrow$ if the derivative exists

Mean (Expectation) and Variance μ_X, σ_X^2

Discrete

$$\mu_X = E(X) = \sum_i x_i f_X(x_i) = \sum_x x\, f_X(x)$$

$$E[g(x)] = \sum_x g(x) f_X(x)$$

$$\sigma_X^2 = V(x) = E[(X - \mu_X)^2] = \sum_x (x - \mu_x)^2 f_X(x)$$

Continuous

$$\mu_X = E(X) = \int_{-\infty}^{\infty} x\, f_X(x)\, dx$$

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

$$\sigma_X^2 = V(x) = E[(X - \mu_X)^2] = \int_{-\infty}^{\infty} (x - \mu_x)^2 f_X(x)\, dx$$

$g(x) = x^k \longrightarrow$ k-th moment of X

① $E(aX + b) = a E(X) + b \longrightarrow$ a, b are constants

② $V(X) = E(X^2) - [E(X)]^2$

③ $V(aX + b) = a^2 V(X)$

Chebyshev's Inequality

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

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

2D RVs, Conditional Probability Distributions

Discrete Random Variables	Continuous Random Variables
joint probability (mass) function $f_{X,Y}(x,y)$ - Must satisfy: ① $f_{X,Y}(x_i,y_j) \geq 0 \quad \forall (x_i,y_j) \in R_{X,Y}$ ② $\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} f_{X,Y}(x_i,y_j) = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} Pr(X=x_i, Y=y_j) = 1$ $Pr((X,Y) \in A) = \underbrace{\sum\sum}_{(x,y)\in A} f_{X,Y}(x,y)$	joint probability density function (p.d.f.) - Must satisfy: ① $f_{X,Y}(x_i,y_j) \geq 0 \quad \forall (x_i,y_j) \in R_{X,Y}$ ② $\iint_{(x,y)\in R_{X,Y}} f_{X,Y}(x,y)\,dx\,dy = 1$ OR $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx\,dy = 1$

Marginal Probability Distributions

$f_X(x) = \sum_y f_{X,Y}(x,y)$ and $f_Y(y) = \sum_x f_{X,Y}(x,y)$	$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy$ and $f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx$

Conditional Distribution

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} \quad \text{and} \quad f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)}$$

$\sum_x f_{X|Y}(x|y) = 1$ and $\sum_y f_{Y|X}(y|x) = 1$	$\int_{-\infty}^{\infty} f_{X|Y}(x|y)\,dx = 1$ and $\int_{-\infty}^{\infty} f_{Y|X}(y|x)\,dy = 1$

Independence

$$f_{X,Y}(x,y) = f_X(x)\,f_Y(y) \quad \forall\, x,y$$

Expectation, Covariance, Correlation

$E(g(X,Y))$

$\sigma_{X,Y}$

$\rho_{X,Y}$

Discrete	Continuous
$E(g(X,Y)) = \sum_x\sum_y g(x,y)\,f_{X,Y}(x,y)$	$E(g(X,Y)) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)\,f_{X,Y}(x,y)\,dx\,dy$
$Cov(X,Y) = E[(X-\mu_X)(Y-\mu_Y)]$ $= \sum_x\sum_y (x-\mu_x)(y-\mu_y)\,f_{X,Y}(x,y)$	$Cov(X,Y) = E[(X-\mu_X)(Y-\mu_Y)]$ $= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-\mu_x)(y-\mu_y)\,f_{X,Y}(x,y)\,dx\,dy$

① $Cov(X,Y) = E(XY) - \mu_X\mu_Y$

② $Cov(aX+b, cY+d) = ac\,Cov(X,Y)$

③ $V(aX+bY) = a^2V(X) + b^2V(Y) + 2ab\,Cov(X,Y)$

$$\rho_{X,Y} = \frac{Cov(X,Y)}{\sqrt{V(X)}\sqrt{V(Y)}}$$

① $-1 \leq \rho_{X,Y} \leq 1$

Special Probability Distributions

Discrete Uniform

$$f_X(x) = \begin{cases} \frac{1}{k} & x = x_1, x_2, \dots, x_k; \\ 0 & \text{otherwise} \end{cases}$$

$$\mu = E(X) = \sum_{\text{all } x} x f_X(x) = \sum_{i=1} x_i \frac{1}{k} = \frac{1}{k}\sum_{i=1}^{k} x_i$$

$$\sigma^2 = V(X) = \sum_{\text{all } x} (x-\mu)^2 f_X(x) = \frac{1}{k}\sum_{i=1}^{k} (x_i - \mu)^2$$

or

$$\sigma^2 = E(X^2) - \mu^2 = \frac{1}{k}\left(\sum_{i=1}^{k} x_i^2\right) - \mu^2$$

Bernoulli

$$f_X(x) = p^x(1-p)^{1-x}, \quad x = 0, 1;$$

$$\mu = E(X) = p$$

$$\sigma^2 = V(X) = p(1-p) = pq$$

Binomial $X \sim B(n,p)$

$$Pr(X=x) = f_X(x) = \binom{n}{x} p^x (1-p)^{n-x} = \binom{n}{x} p^x q^{n-x},$$

$$x = 0, 1, \dots, n, \quad 0 < p < 1, \quad q = 1-p$$

$$\mu = E(X) = np$$

$$\sigma^2 = V(X) = np(1-p) = npq$$

Negative Binomial $X \sim NB(k,p)$

$$Pr(X=x) = f_X(x) = \binom{x-1}{k-1} p^k q^{x-k}$$

$$x = k, k+1, k+2, \dots$$

$$E(X) = \frac{k}{p}$$

$$Var(X) = \frac{(1-p)k}{p^2}$$

$$X \sim NB(1,p) \rightarrow X \sim Geo(p)$$

Continuous Uniform $X \sim U(a,b)$

$$f_X(x) = \begin{cases} \frac{1}{b-a} & a \le x \le b; \\ 0 & \text{otherwise.} \end{cases}$$

$$E(X) = \frac{a+b}{2}$$

$$V(X) = \frac{1}{12}(b-a)^2$$

Poisson $X \sim Po(\lambda)$

$$f_X(x) = Pr(X=x) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x = 0, 1, 2, 3, \dots$$

$$E(X) = \lambda$$

$$V(X) = \lambda$$

Binomial → Poisson

$$X \sim B(n,p) \rightarrow X \overset{\text{approx.}}{\sim} Po(np)$$

$n \rightarrow \infty$ and $p \rightarrow 0$ s.t. $\lambda = np$

$$\lim_{\substack{p \to 0 \\ n \to \infty}} Pr(X=x) = \frac{e^{-np}(np)^x}{x!}$$

Exponential $X \sim exp(\alpha)$

$$f_X(x) = \begin{cases} \alpha e^{-\alpha x} & x > 0 \\ 0 & \text{otherwise} \end{cases}$$

$$E(X) = \frac{1}{\alpha}$$

$$V(X) = \frac{1}{\alpha^2}$$

No memory property

$$Pr(X > s+t \mid X > s) = Pr(X > t)$$

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

$$f_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma}\right), \quad -\infty < x < \infty,$$

$$E(X) = \mu$$

$$V(X) = \sigma^2$$

Standardisation

$$Z = \frac{(X-\mu)}{\sigma} \sim N(0,1)$$

$$E(Z) = 0$$

$$V(Z) = 1$$

$$Pr(x_1 < X < x_2) = Pr(z_1 < Z < z_2)$$

Binomial → Normal

$$X \sim B(n,p) \rightarrow X \overset{\text{approx.}}{\sim} N(np, npq)$$

$n \rightarrow \infty$ and $p \rightarrow \frac{1}{2}$

$\Rightarrow np > 5$ and $n(1-p) > 5$

Continuity Correction

① $Pr(X = k) \approx Pr(k - \frac{1}{2} < X < k + \frac{1}{2})$

② $Pr(a \le X \le b) \approx Pr(a - \frac{1}{2} < X < b + \frac{1}{2})$

$Pr(a < X \le b) \approx Pr(a + \frac{1}{2} < X < b + \frac{1}{2})$

$Pr(a \le X < b) \approx Pr(a - \frac{1}{2} < X < b - \frac{1}{2})$

$Pr(a < X < b) \approx Pr(a + \frac{1}{2} < X < b - \frac{1}{2})$

③ $Pr(X \le c) = Pr(0 \le X \le c) \approx Pr(-\frac{1}{2} < X < c + \frac{1}{2})$

④ $Pr(X > c) = Pr(c < X \le n) \approx Pr(c + \frac{1}{2} < X < n + \frac{1}{2})$

Sampling, Sampling Distributions

Estimation Based on Normal Distribution