概率论（不完全版）

只记到一元连续型随机变量

参考教材：无

日期：2019.12

More：

Theory of Probability.

Frequency 频率.

Hypothesis 假设.

set 集合.

Universal set : 全集.

Subset : 子集.

Empty / Null set : ϕ.

Union , Intersection, Complement

并 交 补($\overline{A}$).

Distributive laws : $A\cup(B\cap C)=(A\cup B)\cap(A\cup C)$.

$A\cap(B\cup C)=(A\cap B)\cup(A\cap C)$.

DeMorgan's laws $\overline{A\cup B}=\overline{A}\cap\overline{B}$.

$\overline{A\cap B}=\overline{A}\cup\overline{B}$.

Venn diagram 维恩图.

Disjoint / Mutually exclusive :

$A\cap B=\phi$.

Experiment : the process by which an ~~obversation~~ observation is made.

Outcomes : Events. A, B, C…

Compound event : 复合的

Simple event :

sample point : 样本点. $a_1, a_2, a_3\cdots$

Sample space : 样本空间 : all possible outcomes.

Discrete …. : 离散的. S可数或有限.

扫描全能王 创建

Basic axioms of probability.

$P(A)$

(1) : $P(A) \geq 0$. (2) $P(S) = 1$. (3) $A_1, A_2, \cdots$ mutually exclusive, then

$P(A_1 \cup A_2 \cup \cdots) = \sum_{i=1} P(A_i)$. 可列可加.

Equally-Likely Outcomes Model 等概模型.

Laplace.

N个 outcomes 等概率发生.

Counting Rules.

$$P(A) = \frac{\#\text{ of simple events in } A}{\#\text{ of } \cdots\ \cdots \text{ in } S}.$$

Rule 1 : $m \times n$.

3×2.

2×3

Rule 2 : Permutations.

$$P_r^n = n(n-1)\cdots(n-r+1) = \frac{n!}{(n-r)!}.$$

Rule 3 : Combinations.

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

$$C_r^n + C_{r+1}^n = C_{r+1}^{n+1}.$$

Rule 4 : Partitions.

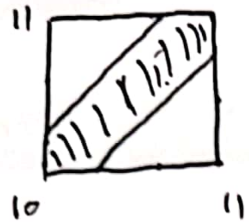
n个人分为k组，每组有$n_1, \cdots n_k$人. $\sum_{i=1}^{k} n_i = n$.

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

扫描全能王 创建

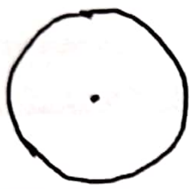
几何概型.

eg. 10点 ~ 11点 碰面, 则到达时刻相差 <20 min 的概率?



$p = \frac{5}{9}$.

贝特朗悖论.



求弦长 > 内接△边长的概率.

2.7. Conditional Probability & Independence.

Definition: B发生时A的概率.

$$p(A|B) = \frac{p(A\cap B)}{p(B)} \quad \text{if } p(B) > 0.$$

Independent Events: $A \perp\!\!\!\perp B \Leftrightarrow p(A|B) = p(A)$

$\Leftrightarrow p(B|A) = p(B)$

$\Leftrightarrow p(AB) = p(A)p(B)$

(用 $p(AB) = p(A|B)p(B)$ 可证)

2.8 Laws of probability.

1. $p(A_1A_2\cdots A_n) = p(A_1)p(A_2|A_1)p(A_3|A_1A_2)\cdots p(A_n|A_1\cdots A_{n-1})$.
2. $p(A\cup B) = p(A) + p(B) - p(AB)$.

   if $A \perp\!\!\!\perp B$ then $p(A\cup B) = p(A) + p(B)$.
3. $p(A) + p(\bar{A}) = 1$.
4. Total Probability: $S = \bigcup_{i=1}^{k} B_i$, $p(B_i) > 0$ $B_i \cap B_j = \phi$

   $p(A) = \sum_{i=1}^{k} p(B_i)p(A|B_i)$.

扫描全能王 创建

5. The Bayes' Rule: $p(B|A) = \dfrac{p(A|B)p(B)}{p(A)}$.

2.10 Random Variables ⇒ 样本空间的实值函数.

Types: Discrete random variable: 离散.

Continuous - - - - : 连续.

eg. $S = \{H, T\}$

$$\underset{\downarrow \atop \text{r.v.}}{X} = \begin{cases} 1 & \text{if } H \\ 0 & \text{if } T \end{cases}$$

Chapter 3. Discrete random variables.

pmf (probability mass function, 概率分布列): a function assigning probabilities to each possible value y of Y.

Properties:
1. $0 \le p(y) \le 1$ for all y
2. if y isn't in the range of Y, $p(y) = 0$
3. $\sum_y p(y) = 1$.

$$P(a \le Y \le b) = \sum_{a \le y \le b} p(y).$$

cdf (cumulative distribution function, 累积分布函数): $F(y) = P(Y \leqq y)$.

Properties:
1. nondecreasing
2. $F(-\infty) = 0$, $F(+\infty) = 1$.

3.3. 期望与方差.

expected value: $\underset{\downarrow \atop \mu}{E(Y)} = \sum_{y \in R} y p(y)$. , R is the range of Y.

扫描全能王 创建

$E(g(Y)) = \sum_y g(y)p(y)$ → 不用求 $g(Y)$ 的 pmf.

Theorem: Y is a r.v.

$$E(g_1(Y) + g_2(Y)) = E(g_1(Y)) + E(g_2(Y)).$$

Extension: X, Y are two r.v s,

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

Variance: $V(Y) = E((Y-\mu_y)^2) = \sum_y (y-\mu_y)^2 p(y) = \delta_y^2$.

Standard deviation: δ_y.

Theorem: $V(Y) = E(Y^2) - (E(Y))^2$. (用期望的线性性质证).

$V(a+bY) = b^2 V(Y)$.

proof: $LHS = E((a+bY - E(a+bY))^2) = E((bY - E(bY))^2) = b^2 E((Y-E(Y))^2)$

$= RHS$.

$\delta(a+bY) = b\,\delta(Y)$.

3.4. Binomial Distribution.

Bernoulli Trial: An experiment with only two possible mutually exclusive outcomes.

Bernoulli distribution: $Y \sim \text{Bernoulli}(p)$.

$$Y = \begin{cases} 1 & p \\ 0 & 1-p \end{cases}, \quad p(Y=y) = p^y(1-p)^{1-y}, \; y=0,1.$$

单次 Bernoulli Trial 基础上的分布.

$E(Y) = p$, $V(Y) = p(1-p)$.

扫描全能王 创建

Binomial Probability Distribution:

Binomial Experiment
- 组成: n 次 Bernoulli Trials
- 每次 $p(S)=p$ 相同
- Trials are independent

Y: number of successes in n trials

Y 服从二项分布. Range Y: $0, 1, \cdots, n$. $Y \sim \text{Binomial}(n; p)$

$$P(Y=y) = \binom{n}{y} p^y (1-p)^{n-y}.$$

$E(Y) = np$; $V(Y) = np(1-p)$. (#).

proof:

$$E(Y) = \sum_{y=0}^{n} y\binom{n}{y} p^y (1-p)^{n-y} = \sum_{y=1} \frac{(n-1)!}{(y-1)!(n-y)!} p^{y-1}(1-p)^{n-y}\cdot(np) = np.$$

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

$$E(Y(Y-1)) = \sum_{y=2} \frac{n!}{(y-2)!(n-y)!} p^y (1-p)^{n-y} = n(n-1)p^2 \sum_{y=2} \frac{(n-2)!}{(y-2)!(n-y)!} p^{y-2}(1-p)^{n-y}$$

$$= p^2(n-1)n$$

$$\therefore V(Y) = p^2 n(n-1) + np - n^2p^2 = np(1-p).$$

Geometric Probability Distribution.

X = # at the first success appears.

$$p(X=x) = (1-p)^{x-1}p, \quad x=1,2,\cdots.$$

$E(Y) = \frac{1}{p}$; $V(Y) = \frac{1-p}{p^2}$

proof: $\frac{1}{p^2} = \frac{1}{[1-(1-p)]^2} = \sum_{y=1}^{\infty} y(1-p)^{y-1}$

So $\sum_{y=1}^{\infty} y(1-p)^{y-1}p = \frac{1}{p}$.

$F(b) = p(X \le b) = 1-(1-p)^b$.

Memoryless property: $p(Y > a+b \mid Y > a) = p(Y > b)$.

无记忆性.

扫描全能王 创建

Hypergeometric Distribution.

N: Population size

r: r items are S.

n: selected.

Y = # of S in the sample of n. $Y \sim HG(n, r, N)$

$$p(Y=y) = \frac{\binom{r}{y}\binom{N-r}{n-y}}{\binom{N}{n}}, \quad \max\{0, n+r-N\} \le y \le \min\{n, r\}.$$

$$E(Y) = n\frac{r}{N}, \quad V(Y) = n\frac{r}{N}\left(1-\frac{r}{N}\right)\frac{N-n}{N-1}.$$

As $N \to \infty$ in such a way that $\frac{r}{N} \to p$, then HG converges to Bin

$$\lim_{N\to\infty} \frac{\binom{r}{y}\binom{N-r}{n-y}}{\binom{N}{n}} = \binom{n}{y} p^y (1-p)^{n-y} \quad \left(\frac{r}{N} \to p\right).$$

Negative Binomial Probability Distribution.

Y = # on which the r^{th} success occurs. $Y \sim NB(r, p)$.

$$p(Y=y) = \binom{y-1}{r-1} p^r (1-p)^{y-r}, \quad y = r, r+1, \ldots.$$

$$E(Y) = \frac{r}{p}; \quad V(Y) = \frac{r(1-p)}{p^2}.$$

Poisson Distribution.

单位时间小概率事件发生的次数.

Poisson Process: 事件独立；强度固定λ；两件及以上几乎不发生.

$$p(Y=y) = \frac{\lambda^y e^{-\lambda}}{y!}, \quad y = 0, 1, 2, \ldots.$$

$$E(Y) = \lambda; \quad V(Y) = \lambda.$$

扫描全能王 创建

近似：$\lim_{n\to\infty}\binom{n}{y}p^y(1-p)^{n-y} = \frac{\lambda^y e^{-\lambda}}{y!}$

small p, large n, $np \to \lambda$.

Moments：

原点矩：k阶：$\mu_k' = E(X^k)$.

$\mu_0' = 1$, $\mu_1' = E(X)$.

中心矩：k阶：$\mu_k = E\{(x-\mu)^k\}$

$\mu_0 = 1$, $\mu_1 = 0$, $\mu_2 = V(X)$.

偏度 Skewness：$\gamma_1 = \frac{\mu_3}{\mu_2^{1.5}}$

$\gamma_1 > 0$：右偏 

$\gamma_1 < 0$：左偏 .

峰度 Kurtosis：$\gamma_2 = \frac{\mu_4}{\mu_2^2} - 3$.

γ_2 大：sharper peak.

性质：$\exists g$, s.t. $\mu_k = g(\mu_0', \mu_1', \dots \mu_k')$.

Moment Generating Function (mgf)

$m(t) = E(e^{tX}) = \sum_x e^{tx} p(x)$.

$\mu_k' = \left.\frac{d^k m(t)}{dt^k}\right|_{t=0} = m^{(k)}(0)$.

mgf 相同 pmf 也相同.

扫描全能王 创建

Bernoulli (p) $\qquad m(t) = pe^t + 1-p.$

Binomial (n,p) $\qquad (pe^t + 1-p)^n$

Geometric (p) $\qquad \dfrac{pe^t}{1-(1-p)e^t}.$

$NB(r,p)$ $\qquad \left(\dfrac{pe^t}{1-(1-p)e^t}\right)^r.$

Poisson (λ) $\qquad e^{\lambda(e^t-1)}.$

~~Chef~~ Chebyshev's Inequality.

$$p(|x-\mu| > k\sigma) < \frac{1}{k^2}.$$

Continuous Random Variables.

CDF : cumulative distribution function

$$F(y) = P(Y \le y).$$

若 $F(y)$ 连续，称 Y 为 连续型 rv.

for any y, $P(Y=y)=0$ (y为 连续型 rv).

Properties: $F(-\infty)=0$; $F(+\infty)=1$

$F(y)$ nondecreasing

pdf: probability density function

$$f(y) = F'(y).$$

$$P(a \le Y \le b) = \int_a^b f(t)\,dt.$$

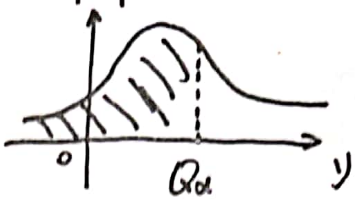
Properties: $f(y) \geqslant 0$

$$\int_{-\infty}^{+\infty} f(y)\,dy = 1.$$

扫描全能王 创建

分位数: the αth quantile of Y, Q_α pdf

$\min Q_\alpha$, s.t $F(Q_\alpha) \geq \alpha$.



$S = \alpha$.

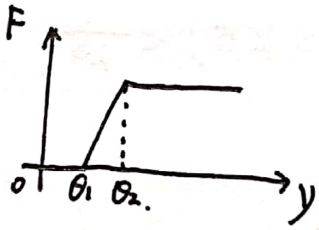
$$E(Y) = \int_{-\infty}^{+\infty} y f(y)\, dy.$$

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

The Uniform Distribution $U(\theta_1, \theta_2)$.

$$f(y) = \begin{cases} \dfrac{1}{\theta_2 - \theta_1} & \theta_1 \le y \le \theta_2 \\ 0 & \text{otherwise.} \end{cases}$$

$$F(y) = \begin{cases} \dfrac{y - \theta_1}{\theta_2 - \theta_1} & \theta_1 \le y \le \theta_2 \\ 1 & y > \theta_2. \end{cases}$$



$E(Y) = \frac{1}{2}(\theta_1 + \theta_2)$

$V(Y) = \frac{1}{12}(\theta_2 - \theta_1)^2$.

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

The Normal Distribution $N(\mu, \delta^2)$

$$f(y) = \frac{1}{\sqrt{2\pi}\,\delta} e^{-\frac{(y-\mu)^2}{2\delta^2}}.$$

Properties: symmetric.
bell-shaped.
centred at μ.

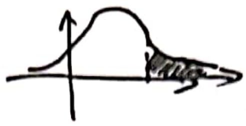
$E(Y) = \mu$, $V(Y) = \delta^2$.

$m(t) = e^{\mu t + \frac{1}{2}\delta^2 t^2}$.

扫描全能王 创建

$Z \sim N(0,1)$.

$$\phi(z) = F(z) = p(Z \le z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt.$$

查表求解：表中给的是  右侧面积，即 $1-\phi(z)$.

$p(-1 \le Z \le 1) = 0.6826$ (1δ)

2δ : 0.9544

3δ : 0.9973

$p(-1.28 \le Z \le 1.28) = 0.8$

0.9 区间 : 1.64

0.95 : 1.96

0.99 : 2.58.

Standardization:

$Y \sim N(\mu, \delta^2)$, then $Z = \frac{Y-\mu}{\delta} \sim N(0,1)$

$p(Y<y) = \phi\left(\frac{y-\mu}{\delta}\right)$.

the αth percentage point of X : x_α. $p(X \geqslant x_\alpha) = \alpha$.

$Q_\alpha = x_{1-\alpha}$.

查表解 Z_α, $X \sim N(\mu, \delta^2)$ then $x_\alpha = \delta Z_\alpha + \mu$.

正态近似二项：

$Y \sim \text{Binomial}(n, p)$, $\mu = np$, $\delta = \sqrt{np(1-p)}$.

$n \to \infty$, $np > 10$, $n(1-p) > 10$,

then $p(Y \le y) \approx p\left(Z \le \frac{y+0.5-\mu}{\delta}\right)$, $Z \sim N(0,1)$.

扫描全能王 创建

Gamma Function

$$\Gamma(\alpha)=\int_0^{+\infty} x^{\alpha-1} e^{-x} dx.$$

$\Gamma(n) = (n-1)!$

$\alpha>1$, $\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1)$.

$\Gamma(\frac{1}{2}) = \sqrt{\pi}$.

The Gamma distribution.

$\alpha>0, \beta>0, y>0.$

$$f(y)=\frac{y^{\alpha-1}e^{-y/\beta}}{\beta^{\alpha}\ \Gamma(\alpha)}. \qquad Y\sim \Gamma(\alpha,\beta)$$

α 控制对称性，β 控制尺度.

$E(Y)=\alpha\beta$, $V(Y)=\alpha\beta^2$

$m(t)=(1-\beta t)^{-\alpha}$.

Exponential distribution.

$Y\sim Exp(\beta)$.

$f(y)=\frac{1}{\beta}e^{-\frac{y}{\beta}}$. $y>0$.

$E(Y)=\beta$, $V(Y)=\beta^2$.

$F(y)=p(Y\le y)=1-e^{-\frac{y}{\beta}}$.

Memoryless property: $p(Y\ge a+b \mid Y\ge a)=p(Y\ge b)$.

场合：独立随机事件的时间间隔；电子产品寿命.

扫描全能王 创建

Chi-square distribution

$\alpha = \frac{\nu}{2}$, $\beta = 2$. ν: degrees of freedom

$Y \sim \chi^2(\nu)$

$$f(y) = \frac{1}{2^{\frac{\nu}{2}}\Gamma(\frac{\nu}{2})} y^{\frac{\nu}{2}-1} e^{-\frac{y}{2}}, \quad y > 0.$$

$E(Y) = \nu$, $V(Y) = 2\nu$.

Beta Function: $$B(\alpha,\beta) = \int_0^1 y^{\alpha-1}(1-y)^{\beta-1}\,dy$$

$$= \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}, \quad \alpha, \beta > 0.$$

$B(\frac{1}{2}, \frac{1}{2}) = \pi$.

Beta Distribution

$$f(y) = \frac{y^{\alpha-1}(1-y)^{\beta-1}}{B(\alpha,\beta)}, \quad 0 \le y \le 1. \ \alpha, \beta > 0$$

若 $\alpha = \beta = 1$, $Y \sim U(0,1)$.

$\alpha = \beta$ 时对称. $\alpha < \beta$ 右偏; $\alpha > \beta$ 左偏.

$E(Y) = \frac{\alpha}{\alpha+\beta}$, $V(Y) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$.

$m(t)$ 无 closed form.

扫描全能王 创建