

Lecture 8

常用分布及其应用-I

Probability and Statistics
Beihang University

8.1 特殊分布总览

下表列出了几个常用的分布

X 的分布	概率或密度函数	$E(e^{tX})$	$E(X)$	$Var(X)$
均匀分布 (a, b)	$f(x) = \frac{1}{b-a} \cdot 1_{(a,b)}(x)$	$\frac{e^{tb} - e^{ta}}{t(b-a)}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
指数分布 $(\beta > 0)$	$f(x) = \beta e^{-\beta x} \cdot 1_{\{x>0\}}(x)$	$\frac{\beta}{\beta - t}$	$\frac{1}{\beta}$	$\frac{1}{\beta^2}$
Gamma 分布 $(n, \lambda > 0)$	$f(x) = \frac{\lambda^n}{\Gamma(n)} x^{n-1} e^{-\lambda x} \cdot 1_{\{x>0\}}(x)$	$\left(\frac{\lambda}{\lambda - t}\right)^n$	$\frac{n}{\lambda}$	$\frac{n}{\lambda^2}$
正太分布 $\mathcal{N}(\mu, \sigma^2)$	$f(x) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}$	$\exp\left\{\mu t + \frac{\sigma^2 t^2}{2}\right\}$	μ	σ^2
二项分布 $(n \in \mathbb{N}, 0 \leq p \leq 1)$	$f(k) = C_n^k p^k (1-p)^{n-k}, k = 0, 1, \dots, n$	$(pe^t + (1-p))^n$	np	$np(1-p)$
Poisson 分布 $(\lambda > 0)$	$f(k) = e^{-\lambda} \frac{\lambda^k}{k!}, k = 0, 1, 2, \dots$	$\exp\{\lambda(e^t - 1)\}$	λ	λ
几何分布 $(0 \leq p \leq 1)$	$f(k) = p(1-p)^{k-1}, k = 1, 2, 3, \dots$	$\frac{pe^t}{1 - (1-p)e^t}$	$\frac{1}{p}$	$\frac{1-p}{p^2}$

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx, \Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1), \alpha > 0, \Gamma(1) = 1.$$

幂级数要点

形如 $\sum_{n=0}^{\infty} a_n (x - x_0)^n$ 的级数称为幂级数. 假设 $a_n \neq 0$, 并且下列极限存在

$$\lim_n \left| \frac{a_{n+1}}{a_n} \right| = \rho$$

那么该幂级数收敛半径为 $\rho^{-1} \in [0, +\infty]$. 在收敛区域内, 幂级数可以 (任意多次) 逐项微分或积分, 并且收敛半径不发生改变. 特别地幂级数 $\sum_{n=0}^{\infty} x^n$ 的收敛域为 $|x| < 1$. 通过逐项微分, 有如下求和公式, $|x| < 1$

//

$$\sum_{n=1}^{\infty} nx^{n-1} = \frac{1}{(1-x)^2}, |x| < 1$$

//

$$\sum_{n=2}^{\infty} n(n-1)x^{n-2} = \frac{2}{(1-x)^3}, |x| < 1$$

矩生成函数（参考内容）

定义 8.1.1 任给随机变量 X , 对 $\forall t \in \mathbb{R}$, 定义

$$\psi(t) = E(e^{tX})$$

$\psi(t)$ 称为 X 的**矩生成函数**.

矩生成函数仅依赖于随机变量的分布. 矩生成函数与随机变量 X 的各阶矩有密切关系.

// 如果存在 $\delta > 0$, 使得 $\psi(t)$ 在某原点领域 $(-\delta, \delta)$ 内收敛, 那么随机变量 X 存在任意阶矩 $\alpha_n = E(X^n), \forall n \in \{0, 1, 2, \dots\}$, 并且

$$\psi(0)^{(n)} = E(X^n)$$

// 如果还知道幂级数 $\sum_n \alpha_k \frac{x^n}{n!}$ 有正收敛半径, 那么 X 的分布由 X 的各阶矩 $\{\alpha_k\}_{k \geq 1}$ 完全确定.

8.2 Bernoulli 分布

定义 8.2.1 随机变量 X 称为参数为 p ($0 \leq p \leq 1$) 的 **Bernoulli 分布**, 如果 X 仅取 0 和 1 两个值, 并且其概率分别为

$$P(X=1)=p, P(X=0)=1-p$$

Bernoulli 随机变量 X 的概率函数可写作

$$f(k|p) = \begin{cases} p^k (1-p)^{1-k}, & k=0, 1 \\ 0, & \text{其他} \end{cases}$$

Bernoulli 随机变量用于描述仅有 0 和 1 两个结果的试验, 试验结果为 1 则称该次**试验成功**. 例如投掷硬币是否出现正面, 临床试验是否成功, 产品是否合格等

定理 8.2.1 随机变量 X 服从参数为 p 的 Bernoulli 分布, 那么

(1) $E(X) = p$;

(2) $Var(X) = p(1 - p)$;

(3) $\psi(t) = E(e^{tX}) = pe^t + (1 - p), t \in \mathbb{R}$

■ (1)

$$E(X) = 1 \cdot p + 0 \cdot (1 - p) = p$$

(2)

$$\begin{aligned} Var(X) &= E(X^2) - (E(X))^2 \\ &= 1^2 \cdot p + 0^2 \cdot (1 - p) - p^2 = p(1 - p) \end{aligned}$$

(3)

$$\begin{aligned} \psi(t) &= E(e^{tX}) \\ &= e^{t \cdot 1} \cdot p + e^{t \cdot 0} \cdot (1 - p) \end{aligned}$$

$$= pe^t + (1 - p)$$



定义 8.2.2 如果随机变量 X_1, X_2, \dots , **独立同分布**, 其中每一个 X_i 是参数为 p 的 Bernoulli 分布. 那么 X_1, X_2, \dots , 被称为具有参数为 p 的 **Bernoulli 实验**. 无穷 Bernoulli 实验也称为 **Bernoulli 过程**.

8.3 二项分布

定义 8.3.1 随机变量 X 称为具有参数 n (正整数) 和 p ($0 \leq p \leq 1$) 的**二项分布**, 如果 X 的概率函数为

$$f(k|n, p) = \begin{cases} C_n^k p^k (1-p)^{n-k}, & k = 0, 1, \dots, n \\ 0, & \text{其他} \end{cases}$$

例题 8.3.1 如果随机变量 X_1, X_2, \dots, X_n 是参数为 p 的 Bernoulli 实验. 那么 $X = \sum_{i=1}^n X_i$ 是参数为 n 和 p 的**二项分布**.

二项分布用于描述 n 次 Bernoulli 试验中的成功次数. 例如投掷硬币 n 次, 正面出现的次数; 对 n 个产品进行检验, 合格产品的数目.

定理 8.3.1 随机变量 X 服从参数 n 和 p 的**二项分布**, 那么

- (1) $E(X) = np$;
- (2) $Var(X) = np(1-p)$;

$$(3) \psi(t) = E(e^{tX}) = (pe^t + (1-p))^n, t \in \mathbb{R}$$

■ 将 X 分解为如例题 (8.3.1) 的独立同分布 Bernoulli 随机变量之和,

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

那么由 Bernoulli 随机变量的性质可得

(1)

$$E(X) = \sum_{i=1}^n EX_i = np$$

(2) X_i 独立, 因此不相关,

$$\text{Var}(X) = \sum_{i=1}^n \text{Var}(X_i) = np(1-p)$$

(3) 由 X_i 独立,

$$\psi(t) = E(e^{tX}) = \prod_{i=1}^n E(e^{tX_i}) = (pe^t + (1-p))^n$$



8.4 几何分布

定义 8.4.1 随机变量 X 称为具有参数 p ($0 \leq p \leq 1$) 的几何分布, 如果 X 的概率函数为

$$f(k|p) = \begin{cases} p(1-p)^{k-1}, & k = 1, 2, \dots \\ 0, & \text{其他} \end{cases}$$

几何分布用于描述 Bernoulli 试验的首次成功时间. 例如投掷硬币直到出现正面所需要的投掷总次数. 因此几何分布也称为 (首次成功) 等待时间的分布.

定理 8.4.1 随机变量 X 服从参数为 p 的几何分布, 那么

$$(1) E(X) = \frac{1}{p};$$

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

$$(3) \psi(t) = E(e^{tX}) = \frac{pe^t}{1-(1-p)e^t}, t < \ln \frac{1}{1-p}.$$

■ (1)

$$\begin{aligned} E(X) &= \sum_{k=1}^{\infty} kp(1-p)^{k-1} = p \sum_{k=1}^{\infty} k(1-p)^{k-1} \\ &= p \cdot \frac{d}{dx} \left(\sum_{k=0}^{\infty} x^k \right) \bigg|_{x=1-p} = p \cdot \frac{1}{p^2} = \frac{1}{p} \end{aligned}$$

(2) 先计算

$$\begin{aligned} E(X^2) &= \sum_{k=1}^{\infty} k^2 p (1-p)^{k-1} \\ &= p \sum_{k=1}^{\infty} k(k-1)(1-p)^{k-1} + p \sum_{k=1}^{\infty} k(1-p)^{k-1} \\ &= p(1-p) \sum_{k=2}^{\infty} k(k-1)(1-p)^{k-2} + \frac{1}{p} \end{aligned}$$

$$\begin{aligned}
&= p(1-p) \frac{d^2}{dx^2} \left(\sum_{k=0}^{\infty} x^k \right) \Big|_{x=1-p} + \frac{1}{p} \\
&= p(1-p) \frac{2}{p^3} + \frac{1}{p} = \frac{2-p}{p^2}
\end{aligned}$$

因此

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

(3) 当 $|(1-p)e^t| < 1$, 即 $t < \ln \frac{1}{1-p}$, $E(e^{tX})$ 存在, 并且

$$\begin{aligned}
&E(e^{tX}) \\
&= \sum_{k=1}^{\infty} e^{tk} p (1-p)^{k-1} \\
&= p e^t \sum_{k=1}^{\infty} [(1-p)e^t]^{k-1}
\end{aligned}$$

$$= \frac{pe^t}{1 - (1 - p)e^t}$$



如果随机变量 X 服从参数为 p 的几何分布, 那么前 k 次 Bernoulli 试验都没有成功的概率即是

$$P(X > k) = (1 - p)^k$$

这也可以通过直接计算得到

$$\begin{aligned} P(X > k) &= \sum_{i=k+1} P(X = i) \\ &= \sum_{i=k+1} p(1 - p)^{i-1} \\ &= p \sum_{i=k+1} (1 - p)^{i-1} \end{aligned}$$

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

定理 8.4.2 (几何分布无记忆性) 随机变量 X 服从参数为 p 的几何分布. 整数 $k \geq 1$. 那么对任意整数 $t \geq 1$ 都有

$$P(X = k + t | X > k) = P(X = t)$$

注记 8.4.1 这一定理表明在已知前 k 次 Bernoulli 试验都没有成功的条件下, 从第 $k + 1$ 次试验起直至成功的等待时间与 (从第 1 次试验开始到) 首次成功的等待时间分布相同, 并且这一分布与前面没有成功的 Bernoulli 试验次数无关, 因此称为无记忆性. 可以证明几何分布是唯一具有无记忆性的离散分布.

■ 由于 $t \geq 1$, $\{X = k + t\} \subset \{X > k\}$, 因此

$$P(X = k + t, X > k)$$

$$\begin{aligned}
 &= P(X = k + t) \\
 &= p(1 - p)^{k+t-1}
 \end{aligned}$$

因此

$$\begin{aligned}
 &P(X = k + t | X > k) \\
 &= \frac{p(1 - p)^{k+t-1}}{(1 - p)^k} \\
 &= p(1 - p)^{t-1} = P(X = t)
 \end{aligned}$$



8.5 Poisson 分布

某银行估计客户在一段时间 $[0, 1]$ 内到来数量的分布情况 (亦可考虑一般的时间区间 $[0, t]$), 为此银行假设客户到来是相互独立的, 并且假设客户到来的速率为 λ (客户数/单位时间). 他们将 $[0, 1]$ 等分为 n 份, 认为每个小时间区间内客户到来的概率是 $p = \lambda/n$, 然后客户数量 X 就近似地用服从参数 n 和 p 的二项分布来解释. 于是当 n 充分大, (从而 p 变得很小)

$$P(X=0) = (1-p)^n = \left(1 - \frac{\lambda}{n}\right)^n \approx e^{-\lambda},$$

$$\begin{aligned}\frac{P(X=k+1)}{P(X=k)} &= \frac{C_n^{k+1} p^{k+1} (1-p)^{n-k-1}}{C_n^k p^k (1-p)^{n-k}} = \frac{\frac{n!}{(k+1)!(n-k-1)!}}{\frac{n!}{k!(n-k)!}} \frac{p}{1-p} \\ &= \frac{(n-k)p}{(k+1)(1-p)} = \frac{\lambda - kp}{(k+1)(1-p)} \approx \frac{\lambda}{k+1},\end{aligned}$$

因此

$$P(X=1) \approx \lambda \cdot e^{-\lambda},$$

$$P(X=2) \approx \frac{\lambda^2}{2} \cdot e^{-\lambda},$$

$$P(X=3) \approx \frac{\lambda^3}{6} \cdot e^{-\lambda},$$

一般地

$$P(X=k) \approx \frac{\lambda^k}{k!} \cdot e^{-\lambda},$$

容易看出

$$\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} = 1$$

因此在 n 充分大的极限情况下, 二项分布近似到一个新的分布, 我们将这个新的分布称为 **Poisson 分布**.

定义 8.5.1 $\lambda > 0$. 随机变量 X 称为具有参数 λ 的 **Poisson 分布**, 如果 X 的概率函数为

$$f(k|\lambda) = \begin{cases} e^{-\lambda} \frac{\lambda^k}{k!}, & k = 0, 1, 2, \dots \\ 0, & \text{其他} \end{cases}$$

定理 8.5.1 随机变量 X 服从参数为 $\lambda > 0$ 的 Poisson 分布, 那么

- (1) $E(X) = \lambda$;
- (2) $\text{Var}(X) = \lambda$;
- (3) $\psi(t) = E(e^{tX}) = e^{\lambda(e^t-1)}, t \in \mathbb{R}$.

■ (1)

$$E(X) = \sum_{k=0}^{\infty} k \cdot e^{-\lambda} \frac{\lambda^k}{k!} = \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda$$

(2) 先计算

$$\sum_{k=0}^{\infty} k(k-1) \cdot e^{-\lambda} \frac{\lambda^k}{k!} = \lambda^2 e^{-\lambda} \sum_{k=2}^{\infty} \frac{\lambda^{k-2}}{(k-2)!} = \lambda^2$$

于是

$$E(X^2) = \sum_{k=0}^{\infty} (k(k-1) + k) \cdot e^{-\lambda} \frac{\lambda^k}{k!} = \lambda^2 + \lambda$$

因此 $Var(X) = \lambda$.

(3)

$$E(e^{tX}) = \sum_{k=0}^{\infty} e^{tk} \cdot e^{-\lambda} \frac{\lambda^k}{k!} = \sum_{k=0}^{\infty} e^{-\lambda} \frac{(e^t \lambda)^k}{k!} = e^{\lambda(e^t - 1)}$$



定理 8.5.2 随机变量 X_1, \dots, X_n 相互独立, X_k 服从参数为 λ_k ($k = 1, 2, \dots, n$) 的 Poisson 分布, 那么 $X_1 + \dots + X_n$ 服从参数为 $\lambda_1 + \dots + \lambda_n$ 的 Poisson 分布.

二项分布与 Poisson 分布的关系

定理 8.5.3 假设 n 为正整数, $0 < p < 1$, 用 $f(k|n, p)$ 表示参数为 n 和 p 的二项分布, $f(k|\lambda)$ 表示参数为 λ 的 Poisson 分布. 如果 $0 < p_n < 1$, $\lim_{n \rightarrow \infty} np_n = \lambda$. 那么对 $\forall k = 0, 1, 2, \dots$,

$$\lim_{n \rightarrow \infty} f(k|n, p_n) = f(k|\lambda).$$

■ 由假设可知 $\lim_{n \rightarrow \infty} p_n = 0$. 对任意给定的 k , 当 $n \rightarrow \infty$ 时有

$$f(k|n, p_n) = \frac{n(n-1) \cdots (n-k+1)}{k!} p_n^k (1-p_n)^{n-k}$$

$$\begin{aligned}
&= \frac{(np_n)^k}{k!} \cdot \frac{n(n-1) \cdots (n-k+1)}{n^k} \cdot (1-p_n)^{n-k} \\
&\rightarrow \frac{\lambda^k}{k!} \cdot 1 \cdot e^{-\lambda+0} \\
&= e^{-\lambda} \frac{\lambda^k}{k!} = f(k|\lambda)
\end{aligned}$$

即在给定条件下二项分布的概率函数收敛到 Poisson 分布的概率函数

例题 8.5.1 一地区患某疾病人数大约占到 1%. 现从该地区随机抽取 200 人, 其中至少有 4 个人患有这种疾病的概率为多大?

■ 我们假设在这 200 人中, 患病人数 X 的真实分布为参数为 200 和 0.01 的二项分布. 这一分布可以由参数为 $\lambda = 200 \cdot 0.01 = 2$ 的 Poisson 分布来逼近, 由 Poisson 分布的概率函数表 (如所附图表) 可知

$$P(X \geq 4) \approx 0.1420.$$



部分 Poisson 分布概率表如下所示 (来自 www)

x	λ									
	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0
0	0.3329	0.3012	0.2725	0.2466	0.2231	0.2019	0.1827	0.1653	0.1496	0.1353
1	0.3662	0.3614	0.3543	0.3452	0.3347	0.3230	0.3106	0.2975	0.2842	0.2707
2	0.2014	0.2169	0.2303	0.2417	0.2510	0.2584	0.2640	0.2678	0.2700	0.2707
3	0.0738	0.0867	0.0998	0.1128	0.1255	0.1378	0.1496	0.1607	0.1710	0.1804
4	0.0203	0.0260	0.0324	0.0395	0.0471	0.0551	0.0636	0.0723	0.0812	0.0902
5	0.0045	0.0062	0.0084	0.0111	0.0141	0.0176	0.0216	0.0260	0.0309	0.0361
6	0.0008	0.0012	0.0018	0.0026	0.0035	0.0047	0.0061	0.0078	0.0098	0.0120
7	0.0001	0.0002	0.0003	0.0005	0.0008	0.0011	0.0015	0.0020	0.0027	0.0034
8	0.0000	0.0000	0.0001	0.0001	0.0001	0.0002	0.0003	0.0005	0.0006	0.0009
9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001	0.0001	0.0002

8.6 Gamma 分布

定义 8.6.1 假设 $\alpha > 0, \beta > 0$. 随机变量 X 称为具有参数 α 和 β 的 **Gamma 分布**, 如果 X 的概率函数为

$$f(x|\alpha, \beta) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

其中

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

称为 Gamma 函数.

当 $\alpha =$ 正整数 $n, f(x|n, \beta)$ 常常用于排队系统的建模, 例如已知客户到来时间分布的情况下, 银行的服务速率的分布, 即在单位时间内服务的客户数量.

习题 8.6.1 试计算 Gamma 分布的期望, 方差, 与生成函数.

8.7 指数分布

回到 Poisson 分布一节银行客户到来数量问题. Poisson 分布已经给出了客户到来的数量的分布, 现在银行还想知道客户到来的时间间隔分布情况. 由于客户到来是相互独立的, 因此各个客户到来之间的时间间隔同分布, 我们只需知道首个客户的到来时间 T 的分布情况即可. 为此将时间数量化, 假设客户只能在时间点 $0, \delta, 2\delta, \dots$ 上以概率 p_δ 到来. 于是任意给定时刻 t 都包括在某小时间区间内: $n\delta \leq t < (n+1)\delta$, 近似地就有

$$P(T > t) \approx P(T > n\delta) = (1 - p_\delta)^n$$

由期望计算公式

$$ET = \sum_{n=0}^{\infty} P(T > n\delta) \delta = \frac{\delta}{p_\delta}$$

这提示在极限情况 ($\delta \rightarrow 0$) 下, 我们应合理地假设, 存在常数 $\alpha > 0$ 使得 $\delta/p_\delta = \alpha$. 这样 (注意 $n\delta \rightarrow t$)

$$P(T > t) \approx (1 - p_\delta)^n \approx \exp(-np_\delta) = \exp\left(-n\delta \cdot \frac{p_\delta}{\delta}\right) \approx e^{-t/\alpha}$$

容易看到

$$1 - P(T > t) \approx 1 - e^{-t/\alpha}$$

是一个分布函数, 这一极限情形下的分布就称为指数分布.

定义 8.7.1 $\beta > 0$. 随机变量 T 称为具有参数 β 的**指数分布**, 如果 T 的密度函数为

$$f(t|\beta) = \begin{cases} \beta e^{-\beta t}, & t > 0 \\ 0, & t \leq 0 \end{cases}$$

指数分布可视为 Gamma 分布的一个特殊情形.

定理 8.7.1 (指数分布无记忆性) 随机变量 X 服从参数为 β 的指数分布. 那么对任意 $t > 0$ 都有

$$P(X \geq h + t | X \geq h) = P(X \geq t)$$

■ 对任意 $h > 0$

$$P(X \geq h) = \int_h^{\infty} \beta e^{-\beta t} dt = e^{-\beta h}$$

因此

$$P(X \geq h + t | X \geq h) = \frac{P(X \geq h + t)}{P(X \geq h)} = \frac{e^{-\beta(h+t)}}{e^{-\beta h}} = e^{-\beta t}$$

