常见概率分布

王树佳
深圳大学经济学院
wangsj@szu.edu.cn

目录

1 离散型随机变量的分布

1.1 $(0-1)$ 分布

定义 1 设随机变量 X 只可能取 0 与 1 两个值，它的分布律是

$$P(X=k)=p^kq^{1-k}, k=0,1; p+q=1(0<p<1)$$

则称 X 服从$(0-1)$ **分布**或**两点分布**，记为 $X\sim B(x,p)$。
其中 $E(X)=p$，$D(X)=p(1-p)$。

1.2 二项分布

定义 2 在 n 次独立重复试验中，记 X 为事件 A 发生的次数，假设在每次试验中事件 A 发生的概率都是 p，则随机变量 X 的分布称为**二项分布**，记为 $X\sim \textit{Binom}(n,p)$，其概率函数为

$$P(X=k|n,p)=\binom{n}{k}(1-p)^{n-k}, (k=0,1,\dots,n; 0<p<1)$$

其中 $E(X)=np, D(X)=np(1-p)$。

1.3 几何分布

定义 3 在 n 次伯努利试验中，试验 k 次才得到第一次成功的机率，即前 $k-1$ 次皆失败，第 k 次成功的概率。其概率函数为

$$f(x)=P(X=x)=p(1-p)^{x-1}, (x=1,2,\dots)$$

则称 X 的分布为**几何分布**，记为 $X\sim \textit{geom}(p)$。

其数学期望和方差分别为

$$E(X)=\frac{1}{p}, D(X)=\frac{1-p}{p^2}.$$

1.4 负二项分布

定义 4 假设有一组独立的伯努利实验，每次实验都有两种结果“成功”和“失败”，X 为实验成功 r 次之前失败的次数，实验持续到 r 次成功，r 为正整数，其概率函数为

$$f(x)=\binom{r+x-1}{x}p^r(1-p)^x, (x=0,1,2,\dots)$$

则称 X 的分布为**负二项分布**，记为 $X\sim \textit{NB}(r,p)$。当 $r=1$ 时，$\textit{NB}(1,p)=\textit{Geom}(p)$。

其数学期望和方差为

$$E(X)=\frac{r(1-p)}{p}, D(X)=\frac{r(1-p)}{p^2}.$$

1.5 超几何分布

定义 5 产品抽样检查中经常遇到一类实际问题，假定在 N 件产品中有 M 件不合格品，即不合格率 $p=M/N$。在产品中随机抽 n 件做检查，发现 k 件不合格品的概率为

$$PX=k=\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{k}}, k\in Z, max\{0,n-N+M\}\le k\le min\{n,M\}$$

则称 X 的分布为**超几何分布**，记为 $X\sim H(N,M,n)$。

其数学期望和方差为

$$\begin{aligned} E(X) &= \frac{nM}{N}, \\ D(X) &= \frac{nM}{N}(1-\frac{M}{N})(\frac{N-n}{N-1}). \end{aligned}$$

1.6 泊松分布

定义 6 设随机变量 X 所有可能取的值为 $0, 1, 2, \ldots$，参数 λ 是单位时间 $[0, 1]$ 内随机事件的平均发生次数，其概率函数为

$$P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}, (x = 0, 1, 2, \ldots)$$

则称 X 服从参数为 λ 的**泊松分布**，记为 $X \sim Posi(\lambda)$。
其中 $E(X) = \lambda$，$D(X) = \lambda$。

2 连续型随机变量的分布

2.1 均匀分布

定义 7 如果随机变量 X 的密度函数 *(pdf)* 为

$$X \sim f(x|a, b) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0, & otherwise \end{cases}$$

则 X 的分布称为 (a, b) 区间的**均匀分布**，记为 $X \sim Unif(a, b)$。

其数学期望和方差分别为

$$\begin{aligned} E(X) &= \frac{a+b}{2}, \\ D(X) &= \frac{(b-a)^2}{12}. \end{aligned}$$

2.2 正态分布

定义 8 若连续型随机变量 x 的概率密度 (pdf) 为

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp(-\frac{(\mathrm{x}-\mu)^2}{2\sigma^2})$$

其中 $\mu, \sigma (\sigma > 0)$ 为常数，则称 X 服从参数为 μ, σ 的**正态分布**，记为 $X \sim N(\mu, \sigma^2)$。
其中 $E(X) = \mu$，$D(X) = \sigma^2$。
当 $\mu = 0, \sigma = 1$ 时，正态分布就成了**标准正态分布**，记为 $X \sim N(0, 1)$。

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{\mathrm{x}^2}{2})$$

2.3 指数分布

定义 9 设连续型随机变量 x 的概率密度 (pdf) 为

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

其中 $\lambda > 0$ 为常数，则称为 X 服从参数 λ 的**指数分布**，记为 $X \sim Exp(\lambda)$。

其数学期望和方差分别为

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

\end{eqnarray*}

2.4 伽马 (Γ) 分布

定义 10 假设随机变量 x 为等到第 α 件事发生所需的等候时间，概率密度函数 (pdf) 为

$$f(x)=\frac{\lambda^{\alpha}}{\Gamma(\alpha)}x^{\alpha-1}e^{-\lambda x},(x>0)$$

则称 X 的分布称为**伽马分布**，分布中的参数 α 为形状参数，λ 为尺度参数，记为 $X\sim\Gamma(\alpha,\lambda)$ 或 $X\sim \mathit{gamma}(\alpha,\lambda)$。

其数学期望和方差分别为

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

2.5 逆伽玛分布

定义 11 如果 $1/X\sim\mathrm{Gamma}(\alpha,\beta)$，则称 X 服从逆伽玛分布，记为 $X\sim\mathrm{IG}(\alpha,\beta)$ 分布，其 *pdf* 为：

$$f(x|\alpha,\beta)=\frac{\beta^{\alpha}}{\Gamma(\alpha)}x^{-(\alpha+1)}\exp(-\beta/x),\ (\alpha>0,\beta>0)$$

其数学期望为

$$\mathrm{E}(X)=\beta/(\alpha-1),\ \alpha>1,$$

其方差为

$$\mathrm{D}(X)=\frac{\alpha^2}{(\alpha-1)^2(\alpha-2)},\ \alpha>2.$$

If $X\sim\mathrm{IG}(\alpha,\beta)$*, then* $1/X\sim\mathrm{Gamma}(\alpha,\beta)$,

If $X\sim\mathrm{IG}(\nu/2,\nu s^2/2)$*, then* $\nu s^2/X\sim\chi^2(\nu)$。

2.6 贝塔 (β) 分布

定义 12 一组定义在 $(0,1)$ 区间的连续概率分布，有两个参数 $\alpha,\beta>0$，其概率密度函数 (pdf) 为

$$\begin{aligned}f(x) &= \frac{1}{B(\alpha,\beta)}x^{\alpha-1}(1-x)^{\beta-1}\\ &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1},(0<x<1)\end{aligned}$$

随机变量 X 服从参数为 α,β 的**贝塔分布**，记为 $X\sim \mathit{Beta}(\alpha,\beta)$。

其数学期望和方差分别为

$$\begin{aligned}E(X) &= \frac{\alpha}{\alpha+\beta},\\ D(X) &= \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}.\end{aligned}$$

2.7 柯西分布

定义 13 柯西分布是一个数学期望不存在的连续型概率分布，其概率密度函数 (pdf) 为

$$\begin{aligned}f(x) &= \frac{1}{\beta\pi}[1+(\frac{x-m}{\beta})^2]^{-1}\\ &= \frac{1}{\pi}\frac{1}{\lambda^2+(x-a)^2},(-\infty<x<\infty)\end{aligned}$$

式中 m 为定义峰值位置的位置参数，β 为尺度参数，记为 $X\sim \mathit{cauchy}(\mathrm{location}=\mathrm{m},\mathrm{scale}=\beta)$。

其数学期望和方差不存在。

2.8 对数正态分布

定义 14 一个随机变量的对数服从正态分布，则该随机变量服从**对数正态分布**，设 X 是取值为正数的连续随机变量，若 $lnX \sim N(\mu, \sigma^2)$，$X$ 的概率密度函数 *(pdf)* 为

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp[\frac{-(\mathrm{lnx} - \mu)^2}{2\sigma^2}], (\mathrm{x} > 0)$$

记为 $X \sim Inorm(\mu, \sigma^2)$。

其数学期望和方差分别为

$$\begin{aligned} E(X) &= e^{\mu + \frac{\sigma^2}{2}}, \\ D(X) &= (e^{\sigma^2} - 1)e^{2\mu + \sigma^2}. \end{aligned}$$

2.9 韦布尔分布/韦伯分布

定义 15 **韦伯分布**是一个连续型的概率分布，是可靠性分析和寿命检验的理论基础，其概率密度函数 *(pdf)* 为

$$f(x) = \frac{\alpha}{\beta}(\frac{x}{\beta})^{\alpha-1} \exp(\frac{\mathrm{x}}{\beta})^{\alpha}, (\mathrm{x} > 0)$$

其中，α 为比例参数，β 是形状参数，记为 $X \sim Weibull(\alpha, \beta)$。

其数学期望和方差分别为

$$\begin{aligned} E(X) &= \alpha\Gamma(1 + \frac{1}{\beta}), \\ D(X) &= \alpha^2[\Gamma(1 + \frac{2}{\beta}) - \Gamma(1 + \frac{1}{\beta})^2]. \end{aligned}$$

2.10 t 分布

定义 16 设 $X \sim N(0,1), Y \sim \chi^2(n)$。且 X, Y 独立，则称随机变量 $t = \frac{X}{\sqrt{Y/n}}$ 服从自由度为 n 的 t 分布，记为 $t \sim t(n)$。t **分布**又称**学生氏分布**，$t(n)$ 分布的概率密度函数 *(pdf)* 为

$$h(t) = \frac{\Gamma[(n+1)/2]}{\sqrt{\pi n}\Gamma(n+2)}(1 + \frac{t^2}{n})^{-(n+1)/2}, (-\infty < x < \infty)$$

其数学期望和方差分别为

$$E(X) = 0, D(X) = \frac{n}{n-2}$$

2.11 一般 t 分布

定义 17 一个随机变量称为 $T \sim t_\nu(\mu, \sigma^2)$, 如果 $T = \mu + \sigma Z$，其中 Z 为标准 t- 分布，其 *pdf* 为

$$f(t|\nu, \mu, \sigma^2) = \frac{\Gamma[(\nu+1)/2]}{(\sigma^2\nu\pi)^{1/2}\Gamma(\nu/2)} \left[1 + \frac{(t-\mu)^2}{\nu\sigma^2}\right]^{-(\nu+1)/2}$$

其中 ν 为自由度，μ 为位置参数，σ 为尺度参数。

其均值为 $E(T) = \mu(\nu > 1)$，其方差为

$$Var(T) = \frac{\nu\sigma^2}{\nu - 2}\ (\nu > 2).$$

如果 $\mu = 0, \sigma^2 = 1$，则为标准 t-分布：$T \sim t(\nu)$。

2.12 多元 t 分布

定义 18 称 p 维随机向量 $\mathbf{y} \sim t_\nu(\boldsymbol{\mu}, \Sigma)$ 或 $t(\boldsymbol{\mu}, \Sigma, \nu)$，如果

$$f(\mathbf{y}) = c|\Sigma|^{-\frac{1}{2}} \left(1 + \frac{1}{\nu}(\mathbf{y} - \boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{y} - \boldsymbol{\mu})\right)^{-(\nu+p)/2}$$

其中 ν 自由度，$\boldsymbol{\mu}$ 为位置参数，Σ 为尺度参数（正定矩阵）。

如果 $\nu = 1$，则为哥西 *(Cauchy)* 分布。

其数学期望为 $E(Y) = \mu, (\nu > 1)$，其方差为

$$D(Y) = \frac{\nu}{\nu - 2}\Sigma, (\nu > 2).$$

2.13 卡方 (χ^2) 分布

定义 19 设 $X_1, X_2, \ldots, X_n$ 是来自总体 $N(0, 1)$ 的样本, 则称统计量 $\chi^2 = X_1^2 + X_2^2 + \cdots + X_n^2$ 服从自由度为 n 的**卡方分布**，记为 $\chi^2 \sim \chi^2(n)$。自由度是指包含独立变量的个数。其概率密度函数 *(pdf)* 为

$$f(y) = \begin{cases} \frac{1}{2^{n/2}\Gamma(n/2)} y^{n/2-1} e^{-y/2}, & y > 0 \\ 0, & \textit{otherwise} \end{cases}$$

其数学期望和方差分别为

$$E(\chi^2) = n, D(\chi^2) = 2n$$

2.14 F 分布

定义 20 设 $U \sim \chi^2(n_1)$,$V \sim \chi^2(n_2)$。且 U, V 独立, 则称随机变量 $F = \frac{U/n_1}{V/n_2}$ 服从自由度为 (n_1, n_2) 的F **分布**，记为 $F \sim F(n_1, n_2)$, 其概率密度为

$$f(y) = \begin{cases} \frac{\Gamma[(n_1+n_2)/2](n_1/n_2)^{n_1/2} y^{(n_1/2)^{-1}}}{\Gamma(n_1/2)\Gamma(n_2/2)[1+(n_1 y/n_2)]^{(n_1+n_2)/2}} & y > 0 \\ 0, & \textit{otherwise} \end{cases}$$

其数学期望和方差分别为

$$\begin{aligned} E(X) &= \frac{n_2}{n_2 - 2}, \\ D(X) &= \frac{2n_2^2(n_1 + n_2 - 2)}{n_1(n_2 - 2)^2(n_2 - 4)}, (n_2 > 4). \end{aligned}$$

2.15 拉普拉斯分布

定义 21 如果随机变量 X 的密度函数 *(pdf)* 为

$$f(x) = \frac{1}{2\lambda}\exp(-\frac{|\mathrm{x} - \mu|}{\lambda})$$

其中 λ, μ 为常数，且 $\lambda > 0$, 则称 X 服从参数为 μ（位置参数）,λ（尺度参数）的**拉普拉斯分布**，记为 $X \sim La(\mu, \lambda)$。
其数学期望和方差分别为 $E(X) = \mu, D(X) = 2\lambda^2$.

2.16 瑞利分布

定义 22 如果随机变量 X 的密度函数 *(pdf)* 为

$$X \sim f(x) = \begin{cases} \frac{x}{\sigma^2}\exp - \frac{\mathrm{x}^2}{(2\sigma^2)}, & x > 0 \\ 0, & \textit{otherwise} \end{cases}$$

则 X 的分布称为**瑞利分布**。

其数学期望和方差分别为

$$\begin{aligned} E(X) &= \sqrt{\frac{\pi}{2}}\sigma, \\ D(X) &= \frac{4-\pi}{2}\sigma^2. \end{aligned}$$

3 常用多维分布

3.1 多项分布

定义 23 假设进行 n 次独立重复试验，每次试验中有 k 个可能的结果，各个结果发生的概率分别为 $p_1, ..., p_k$ *(*其中 $p_i \geq 0, \sum_{i=1}^k p_i = 1$*)*。记 X_i 为第 i 个结果出现的次数，则称随机向量 $\boldsymbol{X} = (X_1, X_2, \ldots, X_k)^T$ 服从多项分布 *(Multinomial distribution)*，参数为 n 和 $\boldsymbol{p} = (p_1, p_2, \ldots, p_k)^T$，记为 $\boldsymbol{X} \sim \mathrm{Multinom}(n, \boldsymbol{p})$，其联合概率函数 *(pmf)* 为

$$P(X_1 = x_1, \ldots, X_k = x_k) = \binom{n}{x_1 x_x \cdots x_k} p_1^{x_1} p_2^{x_2} \ldots p_k^{x_k}, (x_i \geq 0, \sum_{i=1}^k x_i = n).$$

多项分布 $\mathrm{Multinom}(n, \boldsymbol{p})$ 具有如下性质：

1. $\mathrm{E}(\boldsymbol{X}) = n\boldsymbol{p}$, 即 $\mathrm{E}(X_i) = np_i \, (i = 1, 2, \ldots, k)$
2. $\boldsymbol{X}$ 的协方差矩阵 Σ 是对称矩阵，其对角线元素为
$$\sigma_i^2 = np_i(1 - p_i),$$
非对角线元素为
$$\mathrm{Cov}(X_i, X_j) = -np_i p_j \, (i \neq j).$$
3. $\boldsymbol{X}$ 的相关系数矩阵元素为
$$\mathrm{Corr}(X_i, X_j) = \frac{-np_i p_j}{\sqrt{np_i(1-p_i) \cdot np_j(1-p_j)}} = -\sqrt{\frac{p_i p_j}{q_i q_j}}$$
4. 记 $\boldsymbol{X}_{(k-1)} = (X_1, X_2, \ldots, X_{k-1})^T$，则 $\boldsymbol{X}_{(k-1)} \sim \mathrm{Multinom}(n, \boldsymbol{p}_{(k-1)})$，其中
$$\boldsymbol{p}_{(k-1)} = (p_1, \ldots, p_{k-2}, p_{k-1} + p_k)^T.$$

特别, $X_i \sim \mathrm{Bin}(n, p_i)$.

3.2 多元正态分布

定义 24 **多元正态分布**亦称为**多变量高斯分布**。它是一元正态分布向多维的推广。
N 维随机向量 $\boldsymbol{X} = [X_1, \ldots, X_N]^T$ 如果服从多元正态分布，必须满足下面的三个等价条件：

(1) 任何线性组合 $Y = a_1 X_1 + \cdots + a_N X_N$ 服从正态分布。

(2) 存在随机向量 $\boldsymbol{Z} = [Z_1, \ldots, Z_M]^T$（它的每个元素服从独立标准正态分布），向量 $\boldsymbol{\mu} = [\mu_1, \ldots, \mu_N]^T$ 及 $N \times M$ 矩阵 A 满足 $\boldsymbol{X} = A\boldsymbol{Z} + \boldsymbol{\mu}$。

(3) 存在 $\boldsymbol{\mu}$ 和一个对称半正定阵 $\boldsymbol{\Sigma}$ 满足 $\boldsymbol{X}$ 的特征函数：

$$\varphi_X(\boldsymbol{\mu}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \exp(\mathrm{i}\boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{\mu} - \frac{1}{2}\boldsymbol{\mu}^{\mathrm{T}}\Sigma\boldsymbol{\mu})$$

如果一个 p 维随机向量 $\boldsymbol{X}$ 服从**多元正态分布（多元高斯分布）**，其联合概率密度函数为

$$f(\boldsymbol{x}) = \frac{1}{(2\pi)^{p/2}|\boldsymbol{\Sigma}|^{1/2}} exp\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\}, (\boldsymbol{\Sigma} > 0)$$

其中 $\boldsymbol{\mu}$ 是均值向量，$\boldsymbol{\Sigma}$ 是对称正半定矩阵，记为 $\boldsymbol{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$。

其数学期望和方差分别为

$$E(\boldsymbol{X}) = \boldsymbol{\mu}, D(\boldsymbol{X}) = \boldsymbol{\Sigma}.$$

3.3 威希特分布（Wishart 分布）

定义 25 假设 $\boldsymbol{X}_{(\alpha)}(\alpha = 1, 2, ..., n)$ 相互独立，且 $\boldsymbol{X}_{(\alpha)} \sim N_p(0, \boldsymbol{\Sigma})$, 则随机 $p \times p$ 矩阵 $\boldsymbol{W} = \sum_{i=1}^{p} x_i x_i^T$ 服从自由度为 n 的**威希特分布**，记为 $\boldsymbol{W} \sim W_p(n, \boldsymbol{\Sigma})$。其概率密度函数为

$$f(\boldsymbol{W}) = \frac{1}{2^{\frac{np}{2}} |\boldsymbol{\Sigma}|^{\frac{n}{2}} \boldsymbol{\Gamma}_p(\frac{n}{2})} |\boldsymbol{W}|^{\frac{n-p-1}{2}} \exp\{-\frac{1}{2} \mathrm{tr}(\Sigma^{-1} \boldsymbol{W})\}$$

其数学期望和方差分别为

$$E(\boldsymbol{W}) = n\boldsymbol{\Sigma},$$

$$D(w_{ij}) = n(\sigma_{ij}^2 + \sigma_{ii}\sigma_{jj})$$

对于 $n \geq p$ 的矩阵 $\boldsymbol{W}$, 如果 $\boldsymbol{\Sigma}$ 是可逆的, 那么它也是可逆的。当 $p = \boldsymbol{\Sigma} = 1$，那么这个分布是一个自由度为 n 的 χ^2 分布。

3.4 T^2 分布 (Hotelling T^2 分布)

定义 26 假设 $\boldsymbol{W} \sim W_p(n, \boldsymbol{\Sigma})$，$\boldsymbol{X} \sim N_p(0, \boldsymbol{\Sigma})$，$n \geq p$，$\boldsymbol{\Sigma} > 0$, $\boldsymbol{W}$ 与 $\boldsymbol{X}$ 相互独立, 则随机变量

$$T^2 = n\boldsymbol{X}^T \boldsymbol{W}^{-1} \boldsymbol{X}$$

所遵从的分布为第一自由度为 p, 第二自由度为 n 的**T^2 分布**, 记为 $T^2 \sim T^2(p, n)$。T^2 分布可转化为 F 分布, 其关系可以表示为:

$$\frac{n-p+1}{pn} T^2(p, n) = F(p, n-p+1)$$

3.5 Wilks 分布

定义 27 假设 $\boldsymbol{W}_1 \sim \boldsymbol{W}_p(n_1, \boldsymbol{\Sigma})$, $\boldsymbol{W}_2 \sim \boldsymbol{W}_p(n_2, \boldsymbol{\Sigma})$, $\boldsymbol{\Sigma} > 0$, $n_1 > p$,$\boldsymbol{W}_1$ 与 $\boldsymbol{W}_2$ 相互独立, 则

$$\Lambda = \frac{\boldsymbol{W}_1}{\boldsymbol{W}_1 + \boldsymbol{W}_2}$$

服从维度为 p, 第一自由度为 n_1, 第二自由度为 n_2 的 ***Wilks*** **$-\Lambda$ 分布**, 记为 $\Lambda \sim \Lambda(p, n_1, n_2)$。