Chap 4 随机变量的数字特征

1 期望

- **定义(期望)**

$$E(X)=\begin{cases}\sum\limits_i x_i f(x_i)\\ \int\limits_{-\infty}^{\infty} xf(x)dx\end{cases}$$

- **注**
  - 存在 $\Leftrightarrow$ 绝对收敛;
  - **(Lebesque-Stieltjes 积分)** 一般定义: $E(X)=\int\limits_{-\infty}^{\infty} xdF$;
  - 集中趋势的一种刻画;
  - $E((X_1,\cdots,X_n)):=(E(X_1),\cdots,E(X_n))$.
- **性质**
  - $$E(g(X_1,\cdots,X_n))=\begin{cases}\sum\limits_{(x_1,\cdots,x_n)\in\mathbb{R}^n} g(x_1,\cdots,x_n)f(x_1,\cdots,x_n)\\ \int\limits_{\mathbb{R}^n} g(x_1,\cdots,x_n)f(x_1,\cdots,x_n)dx_1\cdots dx_n\end{cases}$$
  - **(线性性质)** $E(aX+bY)=aE(X)+bE(Y)$;
  - 若 $X_1,\cdots,X_n$ 独立, 则 $E(X_1\cdots X_n)=E(X_1)\cdots E(X_n)$.

2 分位数

- **定义(中位数)** X 连续, 若 $P(X\le m)=\dfrac{1}{2}$, 则称 m 为 X 的中位数.
- **注**
  - $F(m)=\dfrac{1}{2}$;
  - $P(X<m)=\dfrac{1}{2}=P(X>m)$;
  - 中位数不一定唯一.
- **定义(中位数)** 若 $P(X<m)\le\dfrac{1}{2}$ 且 $P(X>m)\le\dfrac{1}{2}$, 则称 m 为 X 的中位数.
- **定义(下侧 $\alpha-$分位数)**

  $\forall\alpha\in(0,1)$, 若 $P(X<a)\le\alpha$ 且 $P(X>a)\le 1-\alpha$, 称 a 为 X 的下侧 $\alpha-$分位数.
- **注**
  - 若 X 连续, 则 $P(X<a)=\alpha$;
  - $F^{-1}(\alpha)=\inf\{x\mid F(x)\ge\alpha\}$ 为一个 α 分位数.

- **注**
    - 中位数也是集中趋势的一种刻画;
    - 众数 (方便定义: $f(x)$ 的最大值点).

3 方差

- **定义(方差与标准差)** 给出定义:

$$Var(X) := E((X-E(X))^2) = E(X^2) - E^2(X);$$
$$SD(X) := \sqrt{Var(X)}.$$

- **注** 刻画了数据的集中程度.
- **性质**
    - $Var(c) \equiv 0;$
    - $Var(X+c) \equiv Var(x);$
    - $Var(cX) \equiv c^2 Var(X);$
    - $Var(X+Y) = Var(X) + Var(Y) + 2E((X-E(X))(Y-E(Y))).$
- **注** 定义变异系数 $\frac{\sigma}{\mu}$.

4 协方差与相关系数

- **定义(协方差)**

$$Cov(X,Y) := E((X-\mu_1)(Y-\mu_2)).$$

- **注**
    - $Cov(X,X) = Var(X);$
    - $Cov(X,Y) = Cov(Y,X);$
    - $Cov(X,Y) = E(XY) - \mu_1\mu_2 = E(XY) - E(X)E(Y);$
    - $Cov(aX_1 + bX_2 + c, Y) = aCov(X_1, Y) + bCov(X_2, Y).$
- **定义(协方差矩阵)** 对 $\overline{X} = (X_1, \cdots, X_n), \overline{Y} = (Y_1, \cdots, Y_n)$. 我们有协方差矩阵

$$\begin{aligned} Cov(\overline{X}, \overline{Y}) &= (Cov(\overline{X}_i, \overline{Y}_j))_{n\times n} \\ &= E((\overline{X} - E(\overline{X}))^T(\overline{Y} - E(\overline{Y}))). \end{aligned}$$

- **注** 方差矩阵:

$$\begin{aligned} Var(\overline{X}) &= Cov(\overline{X}, \overline{Y}) \\ &= (Cov(\overline{X}_i, \overline{X}_j))_{n\times n} \\ &= (\sigma_{ij})_{n\times n}. \end{aligned}$$

- **定义(相关系数)**

$$Corr(X,Y)=\frac{Cov(X,Y)}{\sigma_1\sigma_2}=E(\frac{X-\mu_1}{\sigma_1}\cdot\frac{Y-\mu_2}{\sigma_2}).$$

- **定理**
  - 若 X,Y 独立, 则 $Corr(X,Y)=0$, 称为 X,Y 不相关.
  - 联合正态的特殊情况, 不相关可推出独立.
  - $|Corr(X,Y)|\le 1$, 等号成立当且仅当 $\exists a,b$ 使得 $P(Y=aX+b)=1$.
- **证明** 给出引理 **Schwartz 不等式**:

  $$E^2(UV)\le E(U^2)E(V^2).$$

  取等当且仅当 $\exists c\in\mathbb{R}$ 使得 $U=cV$. 取 $U=\dfrac{X-\mu_1}{\sigma_1},V=\dfrac{Y-\mu_2}{\sigma_2}$.
- **注**
  - $\rho:=Corr(X,Y)=\pm1$, 则 $a=\pm\dfrac{\sigma_2}{\sigma_1}$;
  - $\rho:=Corr(X,Y)=0$ (不相关) $\not\Rightarrow$ 独立;

    如 $X\sim N(0,1)$, $Y=X^2$ 不相关但是不独立.
  - 相关系数为线性相关系数.
- **例** $(X_1,X_2)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$, 则

  $$\rho=Corr(X_1,X_2).$$

5 矩

- **定义(矩)** 称

  $$E((X-c)^k)(k=1,2,\cdots)$$

  为 X 关于 c 点的 k 阶矩. 特别地, $c=0$ 对应原点矩, $c=\mu$ 对应中心矩.
- **注**
  - $E(X)=1$ 阶原点矩, $0\equiv 1$ 阶中心矩;
  - $Var(X)=2$ 阶中心矩;
- **定义(偏度系数)**

  $$Skew(X)=\frac{E((X-\mu)^3)}{\sigma^3}=E\left(\left(\frac{X-\mu}{\sigma}\right)^3\right).$$

  称为 3 阶标准矩.
- **注**
  - $0\equiv 1$ 阶标准矩, $1\equiv 2$ 阶标准矩.

- $Skew(X) < 0$ 表示负偏, $Skew(X) > 0$ 表示正偏, 刻画非对称程度;
- 相比于 5 阶及以上的奇数阶矩, 3 阶矩的计算相对简单, 噪声影响较小;
- 不是唯一的刻画偏度的特征数.

- **定义(峰度系数)**

$$Kurt(X) = \frac{E((X-\mu)^4)}{\sigma^4} = E\left(\left(\frac{X-\mu}{\sigma}\right)^4\right).$$

称为 4 阶标准矩.

- **注**
  - 正态分布的峰度 $\equiv 3$, 超额峰度 $:= Kurt(X) - 3$;
  - $Kurt(X) > 3 \leftrightarrow$ 尖峰厚尾;
  - 没有一个数字特征能完美刻画尾部形.

6 矩母函数

- **定义(矩母函数)**

若 $M_X(t) = E(e^{tX})$ 在 $t = 0$ 的某个邻域内存在, 则称 $M_X(t)$ 为 X 的矩母函数. 否则称 X 的矩母函数 **MGF** 不存在.

- **例** $X \sim Exp(\lambda)$.
- **解答**

$$\begin{aligned} M_X(t) &= E(e^{tX}) \\ &= \int_0^{\infty} e^{tx} \lambda e^{-\lambda x} dx \\ &= \frac{\lambda}{\lambda - t},\ t < \lambda. \end{aligned}$$

- **例** $X \sim N(0, 1)$.
- **解答**

$$\begin{aligned} M_X(t) &= E(e^{tX}) \\ &= \frac{1}{\sqrt{2\pi}} \int_0^{\infty} e^{tx} e^{-\frac{1}{2}x^2} dx \\ &= e^{\frac{t^2}{2}},\ t \in \mathbb{R}. \end{aligned}$$

- **性质**
  - $M_X(0) \equiv 1$;
  - $Y = aX + b$, 则 $M_Y(t) = E(e^{t(aX+b)}) = e^{tb} M_X(at)$.
- **例** $Y \sim N(\mu, \sigma^2)$.

- **解答**

$$\begin{aligned} M_Y(t) &= e^{t\mu} M_X(\sigma t) \\ &= e^{\frac{\sigma^2 t^2}{2} + \mu t},\ t \in \mathbb{R}. \end{aligned}$$

- **性质(矩母函数确定矩)**

$$E(X^n) = M_X^{(n)}(0).$$

- **证明**

$$M_X^{(n)}(t) = \sum_{n=0}^{\infty} \frac{M_X^{(n)}(0)}{n!} t^n.$$

又因为

$$\begin{aligned} M_X(t) &= E(e^{tX}) \\ &= E\left(\sum_{n=0}^{\infty} \frac{X^n}{n!} t^n\right) \\ &= \sum_{n=0}^{\infty} \frac{E(X^n)}{n!} t^n. \end{aligned}$$

比较系数即得.

- **例** $X \sim N(0,1)$.
- **解答**

$$\begin{aligned} M_X(t) &= E(e^{tX}) \\ &= e^{\frac{t^2}{2}},\ t \in \mathbb{R} \\ &= \sum_{n=0}^{\infty} \frac{(\frac{t^2}{2})^n}{n!} \\ &= \sum_{n=0}^{\infty} \frac{(2n)!}{2^n n!} \cdot \frac{t^{2n}}{(2n)!} \end{aligned}$$

可得

$$\begin{cases} E(X^{2n+1}) \equiv 0, \\ E(X^{2n}) = \dfrac{(2n)!}{2^n n!}. \end{cases}$$

- **性质(矩母函数确定分布)**

若 $\exists\, a > 0$, 使得 $M_X(t) = M_Y(t), \forall\, t \in (-a, a)$, 则 X, Y 同分布.

- **例** $M_X(t) = \dfrac{1}{4}e^{-t} + \dfrac{1}{2} + \dfrac{1}{8}e^{4t} + \dfrac{1}{8}e^{5t}$.
- **解答** X 离散, 设 $P(X = k) = p_k$, 我们有

$$M_X(t)=E(e^{tX})=\sum_k e^{tk}p_k.$$

可得分布

$$\begin{aligned}P(X=-1)&=\frac{1}{4};\\P(X=0)&=\frac{1}{2};\\P(X=4)&=\frac{1}{8};\\P(X=5)&=\frac{1}{8}.\end{aligned}$$

- **例** $f_1(x)=\frac{1}{\sqrt{2\pi}x}e^{-\frac{(\ln x)^2}{2}},x>0,\ f_2(x)=f_1(x)+f_1(x)\sin(2\pi\ln x).$
- **解答** 注意到

$$\begin{aligned}E(X_2^n)&=E(X_1^n)+\int_0^\infty x^n f_1(x)\sin(2\pi\ln x)dx\\&=0\,(\text{令 } y=\ln x-n)\end{aligned}$$

这是一个同矩不同分布的例子.

- **性质(独立随机变量和的分布)**

若 X, Y 独立, 则 $M_{X+Y}(t)=M_X(t)M_Y(t)$.

- **证明** 注意到

$$\begin{aligned}M_{X+Y}(t)&=E(e^{t(X+Y)})\\&=E(e^{tX}e^{tY})\\&=E(e^{tX})E(e^{tY})\\&=M_X(t)M_Y(t).\end{aligned}$$

- **例** $X_1,X_2,\cdots,X_n$ 独立正态, 则 $X_1+X_2+\cdots+X_n$ 正态.
- **解答** 考察 $X_i\sim N(\mu_i,\sigma_i^2)$, 其中 $i=1,2$. 那么

$$M_{X+Y}(t)=M_X(t)M_Y(t)=e^{\frac{1}{2}(\sigma_1^2+\sigma_2^2)+(\mu_1+\mu_2)t}.$$

进而得到

$$X_1+X_2\sim N(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2).$$

- **注**
  - 若 N 为有限数;
  - 若 N 为随机变量, 与 X_i 独立.
- **注**

- $(X_1, X_2, \cdots, X_n)$ 的 **MGF** 为:

$$M_{(X_1,X_2,\cdots,X_n)}(t_1, t_2, \cdots, t_n) = E(e^{t_1X_1+t_2X_2+\cdots+t_nX_n}).$$

- 特征函数

$$E(e^{itX}), i^2 = -1.$$

7 条件期望

- **定义(条件期望)**

$$E(Y \mid X \in A) = \begin{cases} \sum_i y_i P(Y = y_i \mid X \in A) \\ \int_{-\infty}^{\infty} y f(y \mid X \in A) dy \end{cases}$$

$$E(Y \mid x) = \begin{cases} \sum_i y_i P(Y = y_i \mid X = x) \\ \int_{-\infty}^{\infty} y f_{Y|X}(y \mid x) dy \end{cases}$$

我们称 $E(Y \mid X)$ 为新的随机变量 $h(X)$, 是 Y 对 X 的回归函数.

- **例** $(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$, 则 $E(Y \mid X) = \mu_2 + \rho\frac{\sigma_2}{\sigma_1}(X - \mu_1)$.
- **例** 甲乙两种同类产品, 评价使用寿命为 10 年, 15 年, 市场占有率为 60%, 40%. 随机购买一件产品, 求期望寿命?
- **解答** 为 $10 \times 60\% + 15 \times 40\% = 12$ 年.

若记 X 为产品类型, Y 为产品寿命, 则上式可写成

$$\begin{aligned} E(Y) &= 12 \\ &= 10 \times 60\% + 15 \times 40\% \\ &= E(Y \mid X = 1)P(X = 1) + E(Y \mid X = 2)P(X = 2) \\ &= E[E(Y \mid X)]. \end{aligned}$$

不同取值分层平均并加权.

- **定义(全期望公式)**

$$E(Y) = E[E(Y \mid X)].$$

- **证明** 以连续型为例:

$$\begin{aligned} E(Y \mid x) &= \int_{-\infty}^{\infty} y f_{Y|X}(y \mid x) dy \\ &= \int_{-\infty}^{\infty} y \frac{f(x, y)}{f_X(x)} dy \end{aligned}$$

从而有

$$\begin{aligned}E[E(Y\mid X)]&=\int_{-\infty}^{\infty}E(Y\mid x)f_X(x)dx\\&=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}yf(x,y)dydx\\&=E(Y).\end{aligned}$$

- **注** 一般地, $E[g(X,Y)]=E[E(g(X,Y)\mid X)]$.
- **定理(均方最优预测)**

$$E[(Y-g(X))^2]\geq E[(Y-h(X))^2]=E[(Y-E(Y\mid X))^2]$$

称为均方误差 **MSE** 下的最优预测.

- **证明**

$$E[(Y-c)^2]\geq E[(Y-E(Y))^2].$$

因此

$$E[(Y-g(X))^2\mid X]\geq E[(Y-E(Y\mid X))^2\mid X].$$

两边对 X 取均值, 可得

$$E[(Y-g(X))^2]\geq E[(Y-E(Y\mid X))^2].$$

- **注**
  - $E(Y\mid X)$ 依赖 (X,Y) 的联合分布 (不易获取);
  - 转而求最优线性预测:

$$\min_{a,b}E[(Y-(aX+b))^2]\,(\text{最小二乘法})$$