数字特征

数学期望

离散情况

设X有概率分布 $p_j=P\left(X=x_j\right), j=0,1, \cdots$，如果 $\sum_{j=0}^{\infty}\left|x_j\right| p_j<\infty$，则称 X 的数学期望存在，并且称 $\mathrm{E}(X)=\sum_{j=0}^{\infty} x_j p_j$ 为X或分布 $\left\{p_j\right\}$ 的数学期望。

> X的数学期望EX是其概率分布$\{p_i\}$的重心。

连续情况

假设连续型随机变量X的概率密度为$f(x)$，如果 $\int_{-\infty}^{\infty}|x| f(x) \mathrm{d} x<\infty$，则$X$的数学期望存在，并且称$EX=\int_{-\infty}^{\infty} x f(x) \mathrm{d} x$为$X$或$f(x)$的数学期望。

常用的数学期望

①$X \sim \mathcal{B}(n, p), EX=np$;

②$X \sim \mathcal{P}(\lambda), EX=\lambda$;

③X服从几何分布：$EX=1/p$；（服从负二项分布？求导求和）

④$X \sim \mathcal{E}(\lambda), EX=1/\lambda$。

⑤$X \sim N(0,1), EX^2=1$。（常用的快速计算条件）

数学期望的性质

(1) 若EX存在，且$f(x)$关于$x=c$对称，则$EX=c$;

> 推论：$X \sim N(\mu, \sigma^2), EX=\mu$；$X \sim \mathcal{U}(a, b)=(a+b)/2$。

(2) 如果$P(X \geq 0)=1$，即X是非负的，无论EX是否无穷，都可以直接计算EX。

(3) 设 $\mathrm{E}\left|X_j\right|<\infty(1 \leqslant j \leqslant n), c_0, c_1, \cdots, c_n$ 是常数，则

$$\mathrm{E}\left(c_0+c_1 X_1+c_2 X_2+\cdots+c_n X_n\right)=c_0+c_1 \mathrm{E} X_1+c_2 \mathrm{E} X_2+\cdots+c_n \mathrm{E} X_n$$

(4) 如果X_i相互独立，则$E(\prod X_i)=\prod EX_i$。

关于X的函数的数学期望

(1) $E(g(X))=\int_R g(x) f(x) dx$;

(1.2) $E(g(X))=\sum_{j=1}^{\infty} g(x_j) p_j, X \sim \{p_j | p_j=P(X=x_j), j \geq 1\}$。

(2) $E(h(X, Y))=\iint_{R^2} h(x, y) f(x, y) dx dy$;

(2.2) $\mathrm{E} h(X, Y)=\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} h\left(x_i, y_j\right) p_{ij},(X, Y) \sim\left\{p_{ij} | p_{ij}=P\left(X=x_i, Y=y_j\right), i, j \geqslant 1\right\}$

(3) 若$X \geq 0$，则$EX=\int_{R^+} P(X>x) dx$。（对待非负随机变量计算EX的重要公式）

条件数学期望

(1) $E(Y|X=x)=E(Y|x)=\int_R y f(y|x) dy$;

(2) $E(Y)=\int_R E(Y|x) f_{X=x}(x) dx$;

(3) $EY=E[E(Y|X)]$，在求$E(Y|X)$时，先设$X=x$，在$E(Y|x)$中把x换成X即可得到。

方差

定义

设 $\mu=\mathrm{E}X$，如果 $\mathrm{E}(X-\mu)^2<\infty$，则称 $\sigma^2=\mathrm{E}(X-\mu)^2$为 X 的方差，记做 $\operatorname{Var}(X)$ 或 σ_X^2。称 $\sigma_X=\sqrt{\operatorname{Var}(X)}$ 为 X 的标准差。

> 当 X 有离散分布 $p_j=P\left(X=x_j\right), j\geq 1$ 时，$\operatorname{Var}(X)=\mathrm{E}(X-\mu)^2=\sum\limits_{j=1}^{\infty}\left(x_j-\mu\right)^2 p_j$。
>
> 当 X 有概率密度 $f(x)$ 时，$\operatorname{Var}(X)=\mathrm{E}(X-\mu)^2=\int_{-\infty}^{\infty}(x-\mu)^2 f(x) \mathrm{d} x$.

上面两个计算公式不重要，重要的这个计算公式：$DX=EX^2-(EX)^2$。

常用的方差

①$X\sim\mathcal{B}(n,p), DX=npq;$

②$X\sim\mathcal{P}(\lambda), DX=\lambda;$

③X服从几何分布：$DX=q/p^2$；（服从负二项分布？）

④$X\sim\mathcal{E}(\lambda), DX=1/\lambda^2;$

⑤$X\sim\mathcal{U}(a,b), DX=(b-a)^2/12;$

⑥$X\sim N(\mu,\sigma^2), DX=\sigma^2$。

> 如果要用求导求和等方法直接求EX^2不方便，可以先求$E[X(X-1)]$。

方差的性质

设 a, b, c 是常数, $EX=\mu, \operatorname{Var}(X)<\infty, \mu_j=EX_j, \operatorname{Var}\left(X_j\right)<\infty(1\leqslant j\leqslant n)$，则

(1) $\operatorname{Var}(a+bX)=b^2\operatorname{Var}(X)$;

(2) $\operatorname{Var}(X)=\mathrm{E}(X-\mu)^2<\mathrm{E}(X-c)^2$，只要常数 $c\neq\mu$;

(3) $\operatorname{Var}(X)=0$ 的充分必要条件是 $P(X=\mu)=1$;

(4) 当 $X_1, X_2, \cdots, X_n$ **相互独立时**,

$$\operatorname{Var}\left(\sum_{j=1}^n X_j\right)=\sum_{j=1}^n \operatorname{Var}\left(X_j\right)$$

上面的性质多少都反映了$Var(X)$的定义：偏离$x=\mu$的程度。

> **推论**：$D(\bar{X})=D(\frac{1}{n}\sum X_j)=\frac{1}{n^2}D(\sum X_j)=\frac{1}{n}D(X)$。（在后面参数估计中会用到）
>
> **说明**：测量n次取平均值，方差降低n倍，说明只要测量仪器没有系统误差，测量精度总可以通过多次测量取平均来改进。

标准化随机变量

设 $\sigma^2=\operatorname{Var}(X)<\infty, Y=\frac{X-\mathrm{E}X}{\sigma}$, 则 $\mathrm{E}Y=0, \operatorname{Var}(Y)=\frac{1}{\sigma^2}\operatorname{Var}(X-\mathrm{E}X)=1$.

这时称Y是X的标准化。特别地，当$X\sim N(\mu,\sigma^2), Y\sim N(0,1)$。

协方差和相关系数

定义

为了研究随机变量 X, Y 的关系，引入协方差和相关系数的定义。设 $\sigma_X=\sqrt{\sigma_{XX}}, \sigma_Y=\sqrt{\sigma_{YY}}$ 分别是 X, Y 的标准差。设 $\mu_X=\mathrm{E}X, \mu_Y=\mathrm{E}Y$.

(1) 当 σ_X, σ_Y 存在时，称

$$\mathrm{E}\left[\left(X-\mu_X\right)\left(Y-\mu_Y\right)\right]$$

为随机变量 X, Y 的协方差，记做 $\operatorname{Cov}(X, Y)$ 或 σ_{XY}。当 $\operatorname{Cov}(X, Y)=0$ 时，称 X, Y **不相关。**

> 独立一定不相关，但是不相关不一定独立，这里说的不相关是不线性相关。尽管不线性相关，他们之间可能还有其他的非线性关系。

(2) 当 $0<\sigma_X\sigma_Y<\infty$, 称

$$\rho_{XY}=\frac{\sigma_{XY}}{\sigma_X\sigma_Y}$$

为 X, Y 的相关系数。相关系数 ρ_{XY} 也常用 $\rho(X, Y)$ 表示。

协方差计算的常用公式

$\sigma_{XY}=E(XY)-(EX)(EY)$

协方差矩阵

称随机向量 (X_1,X_2) 的协方差 $\sigma_{ij}=\operatorname{Cov}\left(X_i,X_j\right)$ 构成的矩阵

$$\boldsymbol{\Sigma}=\begin{pmatrix}\sigma_{11} & \sigma_{12}\\ \sigma_{21} & \sigma_{22}\end{pmatrix}$$

为 $\boldsymbol{X}$ 的协方差矩阵。因为 $\sigma_{ij}=\sigma_{ji}$，所以协方差矩阵 $\boldsymbol{\Sigma}$ 是对称矩阵。

正态分布参数计算

设 Y_1,Y_2 独立都服从标准正态分布，$ad-bc\neq 0$ 和

$$\begin{cases}X_1=aY_1+bY_2+\mu_1\\ X_2=cY_1+dY_2+\mu_2\end{cases}$$

则 $\operatorname{Cov}(X,Y)=ac+bd.$

$$(X_1,X_2)\sim N\left(\mu_1,\mu_2;\sigma_1^2,\sigma_2^2;\rho\right)$$

其中

$$\sigma_1^2=a^2+b^2,\sigma_2^2=c^2+d^2,\rho=(ac+bd)/\left(\sigma_1\sigma_2\right).$$

定理：如果 $(X_1,X_2)\sim N\left(\mu_1,\mu_2;\sigma_1^2,\sigma_2^2;\rho\right)$，则 X_1,X_2 独立的充要条件是 X_1,X_2 不相关。

矩

设 X 为随机变量，c 为常数，k 为正整数。则 $E\left[(X-c)^k\right]$ 称为 X 关于 c 点的 k 阶**矩**。

比较重要的有两种情况：

（1）$c=0$，这时 $\alpha_k=E\left(X^k\right)$ 称为 X 的 k 阶原点矩。

（2）$c=E(X)$，这时 $\mu_k=E\left[(X-EX)^k\right]$ 称为 X 的 k 阶中心矩。

一阶原点矩就是期望，一阶中心矩 $\mu_1=0$；二阶中心矩 μ_2 就是 X 的方差 $\operatorname{Var}(X)$。在统计学上，高于四阶的矩极少使用，三、四阶矩有些应用，但也不很多。