# 随机变量的数字特征

## 数字特征

**NOTE:** 数字特征是数，不再是随机量。

## 数学期望

意义：衡量随机变量可能取值的平均水平

### 离散型随机变量的数学期望

- 定义：设 $X$ 为离散型随机变量，其概率分布为：

$$P(X=x_k)=p_k,\ \ k=1,2,\cdots$$

  若无穷级数 $\sum_{k=1}^{+\infty}x_kp_k$ **绝对收敛**，则称其为 $X$ 的数学期望，简称为期望，记作 $E(X)$，即

$$E(X)=\sum_{k=1}^{+\infty}x_kp_k$$

- 常见离散型分布的数学期望：
  1. 0-1 分布：$E(X)=p$
  2. 二项分布 $B(n,p)$： $E(X)=np$
  3. 泊松分布 $P(\lambda)$： $E(X)=\lambda$
  4. 超几何分布 $X\sim(N,M,n)$： $E(X)=nM/N$
  5. 几何分布 $P(X)=q^{k-1}p:E(X)=1/p$

### 一般随机变量的数学期望

- 定义：对随机变量 $X$ ，任给 $\epsilon>0$ ，构造随机变量 $X^*$ ：将 $X$ 的所有可能取值即数轴划分为长度为 $\epsilon$ 的一系列区间。并定义新的随机变量

$$X^*=k\epsilon,\quad iff\ X\in[k\epsilon,(k+1)\epsilon)$$

  通过 $X^*$ 的期望的极限来定义 $X$ 的期望，即

$$E(X)=\lim_{\epsilon\to 0}E(X^*)$$

### 连续型随机变量的数学期望

- 定义：设连续型随机变量 $X$ 的概率密度为 $f(x)$ ，若广义积分 $\int_{-\infty}^{+\infty}xf(x)dx$ 绝对收敛，则称此积分为 $X$ 的数学期望，记作 $E(X)$ ，即

$$E(X)=\int_{-\infty}^{+\infty}xf(x)dx$$

- 常见连续型分布的数学期望：
  1. 均匀分布：$U[a,b]$ ：

$$E(X)=\int_a^b\frac{x}{b-a}dx=\frac{1}{2}(a+b)$$

  2. 指数分布：$E(\lambda)$ ：

$$E(X)=\int_0^{+\infty}x\cdot\lambda e^{-\lambda}x=\frac{1}{\lambda}$$

  3. 正态分布：$N(\mu,\sigma^2)$ ：

$$E(X)=\int_{-\infty}^{+\infty}x\cdot\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx=\mu$$

  4. $\Gamma$ 分布：$\Gamma(\alpha,\beta)$

$$E(X)=\int_0^{+\infty}x\cdot\frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x}=\frac{\alpha}{\beta}$$

## 随机变量函数的数学期望

1. 随机变量函数 $Y=g(X)$ 的数学期望：
   - 若 $X$ 为离散型随机变量，其概率分布为

     $$P(X=x_i)=p_i,\ \ i=1,2,\cdots$$

     若无穷级数 $\sum_{i=1}^{+\infty}g(x_i)p_i$ 绝对收敛，则

     $$E(Y)=\sum_{i=1}^{+\infty}g(x_i)p_i$$

   - 若 $X$ 为连续型随机变量，其概率密度为 $f(x)$，若广义积分

     $$\int_{-\infty}^{+\infty}g(x)f(x)dx$$

     绝对收敛，则

     $$E(Y)=\int_{-\infty}^{+\infty}g(x)f(x)dx$$

2. 二维随机变量 $Z=g(X,Y)$ 的数学期望：
   - 若 $(X,Y)$ 为离散型，其概率分布为:

     $$P(X=x_i,Y=y_i)=p_{ij},\ \ i,j=1,2,\cdots$$

     若级数 $\sum_{i,j=1}^{+\infty}g(x_i,y_i)p_{ij}$ 则

     $$E(Z)=\sum_{i,j=1}^{+\infty}g(x_i,y_i)p_{ij}$$

   - 若 $(X,Y)$ 为连续型，其概率密度为 $f(x,y)$，若广义积分

     $$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}g(x,y)f(x,y)dxdy$$

     绝对收敛，则

     $$E(Z)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}g(x,y)f(x,y)dxdy$$

## 数学期望的性质

1. $E(C)=C$
2. $E(aX)=aE(X)$
3. $E(X+Y)=E(X)+E(Y)$
4. 若 $X$，$Y$ 相互独立，则 $E(XY)=E(X)E(Y)$
5. 若存在数 $a$ 使 $P(X\geq a)=1$，则 $E(X)\geq a$
6. 若存在数 $b$ 使 $P(X\leq b)=1$，则 $E(X)\leq b$
7. Schwarz 不等式：$[E(XY)]^2\leq E(X^2)E(Y^2)$
   - 证明: $E((\lambda X+Y)^2)=\lambda^2E(X^2)+2\lambda E(X)E(Y)+E(Y^2)\geq 0$，然后利用判别式即可得到结论。

# 方差

- 意义：衡量随机变量可能取值偏离其平均水平的平均偏离程度

- 定义：若 $E((X-E(X))^2)$ 存在，则称其为随机变量 $X$ 的方差，记作 $D(X)$ 或 $Var(X)$ ，即

$$D(X)=E((X-E(X))^2)$$

称 $\sigma=\sqrt{D(X)}$ 为 $X$ 的均方差或标准差。

- 计算方差的常用公式

$$D(X)=E(X^2)-E(X)^2$$

## 离散型随机变量的方差

- 定义：若 $X$ 为离散型随机变量，其概率分布为 $P(X=x_k)=p_k,\ \ k=1,2,\cdots$ 则 $X$ 的方差：

$$D(X)=\sum_{k=1}^{+\infty}(x_k-E(X))^2\cdot p_k$$

## 连续型随机变量的方差

- 定义：若 $X$ 为连续型随机变量，其概率密度为 $f(x)$ ，则 $X$ 的方差为

$$D(X)=\int_{-\infty}^{+\infty}(x-E(X))^2\cdot f(x)dx$$

## 常见分布的方差

1. 0-1 分布：

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

2. 二项分布 $B(n,p)$ ：

$$\begin{aligned}&E(X)=np\\&E(X^2)=\sum_{t=0}^{n}t^2C_n^tp^tq^(n-t)=n(n-1)p^2\\&D(X)=E(X^2)-E^2(X)=np(1-p)\end{aligned}$$

3. 泊松分布 $P(\lambda)$ ：

$$\begin{aligned}&E(X)=\lambda\\&E(X^2)=\sum_{n=0}^{+\infty}n^2\frac{\lambda^n}{n!}e^{\lambda}=\lambda^2+\lambda\\&D(X)=E(X^2)-E^2(X)=\lambda\end{aligned}$$

4. 几何分布 $p_k=(1-p)^{k-1}p$ ：

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

5. 均匀分布 $U(a,b)$ ：

$$
\begin{aligned}
&E(X)=\frac{a+b}{2}\\
&E(X^2)=\int_a^b x^2\frac{1}{b-a}dx=\frac{1}{3}(a^2+ab+b^2)\\
&D(X)=E(X^2)-E^2(X)=\frac{(b-a)^2}{12}
\end{aligned}
$$

6. 指数分布 $E(X)=\lambda e^{-\lambda x},\ x>0$：

$$
\begin{aligned}
&E(X)=\frac{1}{\lambda}\\
&E(X^2)=\int_0^{+\infty}x^2\lambda e^{-\lambda x}dx=\frac{2}{\lambda^2}\\
&D(X)=E(X^2)-E^2(X)=\frac{1}{\lambda^2}
\end{aligned}
$$

7. 正态分布 $E(X)=N(\mu,\sigma^2)$：

$$
\begin{aligned}
&E(X)=\mu\\
&D(X)=\sigma^2
\end{aligned}
$$

8. $\Gamma$ 分布：

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

### 方差的性质

设 a, C 为常数

1. $D(C)=0$
2. $D(aX)=a^2D(X)$
3. $D(aX+b)=a^2D(X)$
4. $D(X\pm Y)=D(X)+D(Y)\pm 2E((X-E(X))(Y-E(Y)))$，特别地，若 $X,Y$ 相互独立，则 $D(X+Y)=D(X)+D(Y)$，但逆命题不为真。
5. 若 $X_1,X_2,\cdots,X_n$ 相互独立，$a_1,a_2,\cdots,a_n$ 为常数，则 $D(\sum_{i=1}^n a_iX_i+b)=\sum_{i=1}^n a_i^2D(X_i)$
6. 对任意常数 $C$，$D(X)\le E(X-C)^2$，且当且仅当 $C=E(X)$ 时，等号成立。
7. $D(X)=0\leftrightarrow P(X=E(X))=1$ 即 $X$ 服从单点分布。

**NOTE1:** 切比雪夫不等式：对随机变量 $X$，设$E(X),D(X)$ 存在，则对任意 $\epsilon>0$，有:

$$P(|X-E(X)|>\epsilon)\le\frac{D(X)}{\epsilon^2}$$

**NOTE2:** 不是所有的随机变量都有数学期望和方差；对于某些特殊分布类型，在已知数学期望和方差的情况下，可以完全确定其分布，比如正态分布。

## 分位数，中位数

- 定义：设随机变量 $X$，实数 $a$ 和 $p(0<p<1)$，若下式成立

  $$P(X\le a)\ge p\ge P(X<a)$$

  则称 $a$ 是 $X$ 的 $p$ 分位数，记为 $x_p=a$。0.5 分位数又称 $X$ 的中位数，记为 $med(X)$。

- 性质：
  1. 中位数与数学期望类似，是从“平均水平”角度刻画随机变量取值的特性，中位数的优点在与它总是存在，且不受极端取值的影响，但是不具有可加性。

2. p 分位数一定存在但是不一定唯一。比如分布密度曲线某一段 $f(x)=0$，若 p分位数刚好落在这一区间内，则这一段都是 p分位数。

- 下侧分位数和上侧分位数
  - $x_p$ 称为 $X$ 的下侧分位数，如果

$$\int_{-\infty}^{x_p} f(x)dx = p$$

  - $x_p$ 称为 $X$ 的上侧分位数，如果

$$\int_{x_p}^{+\infty} f(x)dx = p$$

# 协方差

- 用来衡量可能不独立的随机变量之间的 **（线性）关系** 的数字特征。
- 定义：对二维随机变量 $(X,Y)$，若 $E([X-E(X)][Y-E(Y)])$ 存在，称其为二维随机变量 $(X,Y)$ 的协方差，记为

$$cov(X,Y) = E([X-E(X)][Y-E(Y)])$$

## 协方差的性质

1. $cov(X,Y)=cov(Y,X)=E(XY)-E(X)E(Y)$，若协方差为 0，称 $X$，$Y$ 不相关，但是 $X$，$Y$ 不一定相互独立；反之如果 $X$，$Y$ 相互独立，那么 $X$，$Y$ 一定不相关。

2. $cov(X,X)=D(X)$

3. $cov(a_1X+b_1,a_2X+b_2)=a_1a_2cov(X,Y)$

4. $cov(X+Y,Z)=cov(X,Z)+cov(Y,Z)$

5. $D(X+Y)=D(X)+D(Y)+cov(X,Y)$

6. $|cov(X,Y)|^2 \le D(X)D(Y)$
   - 证明：利用数学期望的性质 7 即

$$[E^2((X-E(X))(Y-E(Y)))] < E((X-E(X))^2)E((Y-E(Y))^2)$$

   即可

## 协方差矩阵

- 定义：设 n 维随机变量 $(X_1,X_2,\cdots,X_n)$，记 $\sigma_{ij}=cov(X_i,X_j)$，若 $\sigma_{ij}$ 都存在，则称

$$\Sigma(X_1,X_2,\cdots,X_n) = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_{nn} \end{bmatrix}$$

  为 $(X_1,X_2,\cdots,X_n)$ 的协方差矩阵。
- 性质：
  1. 对称非负性
  2. $\sigma_{ii}=D(X_i)$
  3. $\sigma_{ij}^2 \le \sigma_{ii}\sigma_{jj}$

# 相关系数

- 定义：设二维随机变量 $(X,Y)$，若其协方差存在，且 $D(X)>0,\ D(Y)>0$，则称

$$\rho_{XY} = \frac{cov(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}$$

  为 $X,Y$ 的相关系数。若 $\rho_{XY}=0$，则称 $X,Y$ 不相关。

- 性质：
  1. $|\rho_{XY}| \le 1$
     - 证明：由协方差性质 6 即可。

  2. $|\rho_{XY}| = 1$ 等价于 $X$ 和 $Y$ 有线性关系的概率为 1 。即存在常数 $a, b$ 使得 $P(Y = aX + b) = 1$

  3. $$\begin{aligned}\rho_{XY} = 0 &\Leftrightarrow \text{X,Y 不相关} \\ &\Leftrightarrow cov(X, Y) = 0 \\ &\Leftrightarrow E(XY) = E(X)E(Y) \\ &\Leftrightarrow D(X + Y) = D(X) + D(Y)\end{aligned}$$

  4. $X, Y$ 相互独立可以推出 $X, Y$ 不相关；但是反之不成立。当 $(X, Y)$ 服从二维正态分布时 $X, Y$ 独立等价于 $X, Y$ 不相关。

# 随机变量的矩

- $X$ 的 k 阶原点矩：$E(X^k)$
- $X$ 的 k 阶绝对原点矩：$E(|X|^k)$
- $X$ 的 k 阶中心矩：$E((X - E(X))^k)$
- $X$ 的 k+l 阶混合原点矩：$E(X^k Y^l)$
- $X$ 的 k+l 阶混合中心矩：$E((X - E(X))^k (Y - E(Y))^l)$