目　录

第三章　随机变量的数字特征

教学目的:

1) 理解随机变量的数学期望、方差的概念, 并会运用它们的基本性质计算具体分布的期望、方差.
2) 掌握二项分布、Poisson分布、均匀分布、指数分布、正态分布的数学期望和方差.
3) 会根据随机变量的概率分布计算其函数的数学期望.
4) 理解协方差、相关系数的概念, 掌握它们的性质, 并会利用这些性质进行计算, 了解矩的概念.
5) 理解大数定理与中心极限定理。

在前章中, 我们讨论了随机变量的概率分布, 这种分布是随机变量的概率论性质最完整的刻画. 而随机变量的数字特征是某些由随机变量的分布所决定的常数, 它刻画了随机变量或者说刻画了其分布的某一方面的性质, 这些性质往往是实际应用中人们比较关心的. 例如, 我们在了解某一行业工人的经济状况时, 我们首先关心的恐怕会是其平均收入, 这会给我们一个总体的印象, 而收入的分布状况, 倒不一定是最重要的, 这就是刻画总体平均值的数字特征. 另一类重要的数字特征, 是用来衡量随机变量取值的分散程度. 还拿我们上个例子说明, 如果我们考虑两个行业工人的经济状况, 他们的平均收入大体相近, 但是一个行业收入分配较平均, 即大多数人的收入都在平均值上下不远处, 分散程度就小; 另一个行业则相反, 其收入远离平均值很多, 分散程度就大, 这两者的实际意义当然很不相同. 平均值和分散度是刻画随机变量性质的两类最重要的数字特征. 除了这两者之外, 对于多维变量而言, 还有一类刻画各分量之间关系的数字特征, 较为常用的是协方差和相关系数, 这些我们将在下面的章节详细讨论. 数字特征另一个重要意义在于, 当我们不知道随机变量的确切概率分布, 但是清楚其数字特征的情形下, 我们可以根据这些数字特征推断该随机变量大致的概率性质. 比如某个工厂生产一批灯泡, 我们想了解这批灯泡的质量如何. 我们不知道这批灯泡寿命的确切概率分布, 但是如果我们知道这批灯泡的平均寿命, 知道这批灯泡寿命的分散程度, 那我们就可以大致推断出这批灯泡的质量状况.

§3.1 数学期望(均值)及中位数

§3.1.1 数学期望

数学期望也称均值, 是随机变量的一个最基本的数字特征. 我们先看如下的一个例子

例 3.1.1. *一甲乙两人赌技相同, 各出赌金100元, 约定先胜三局者为胜, 取得全部200元. 现在甲胜2局乙胜1局的情况下中止, 问赌本该如何分?*

解: 如果继续赌下去而不中止, 则甲有3/4的概率取胜, 而乙胜的概率为1/4. 所以, 在甲胜2局乙胜1局的这个情况下, 甲能期望"得到"的数目, 应当确定为

$$200 \times \frac{3}{4} + 0 \times \frac{1}{4} = 150(\text{元}),$$

而乙能"期望"得到的数目, 则为

$$200 \times \frac{1}{4} + 0 \times \frac{3}{4} = 50(\text{元}).$$

如果引进一个随机变量X, X等于在上述局面(甲值2胜乙1胜)之下, 继续赌下去甲的最终所得, 则X有两个可能的值: 200 和0, 其概率分别为3/4和1/4. 而甲的期望所得, 即X的"期望"值, 即等于

$$X\text{的可能值与其概率之积的累加}$$

这就是"数学期望"这个名称的由来. 另一个名称"均值" 形象易懂, 也很常用. 下面我们就给出数学期望(均值)的定义:

对一般的离散型分布, 我们有

定义 3.1.1. *设X为一离散型随机变量, 其分布律为*

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

如果$\sum\limits_{i=1}^{\infty} |x_i| p_i < +\infty$, 则称

$$\sum_{i=1}^{\infty} x_i p_i$$

为随机变量X的数学期望(均值), 用符号EX表示. 若$\sum\limits_{i=1}^{\infty} |x_i| p_i = +\infty$, 则称$X$的数学期望(均值)不存在.

对连续型随机变量, 其数学期望的定义如下

定义 3.1.2. 如果连续型随机变量X具有密度函数$f(x)$, 则当

$$\int_{-\infty}^{\infty}|x|f(x)dx<\infty$$

时, 我们将积分

$$\int_{-\infty}^{\infty}xf(x)dx$$

的值称为X的数学期望, 记作EX. 如果

$$\int_{-\infty}^{\infty}|x|f(x)dx=\infty,$$

则称X的数学期望不存在.

下面求解几种常见分布的数学期望.

1. 二项分布$X\sim B(n,p)$:

$$\begin{aligned} EX &= \sum_{k=0}^{n}k.\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k} \\ &= np.\sum_{i=0}^{n-1}\frac{(n-1)!}{i!(n-1-i)!}p^i(1-p)^{n-1-i} \\ &= np \end{aligned}$$

2. Poisson 分布$X\sim P(\lambda)$:

$$EX=\lambda$$

3. 正态分布$X\sim N(\mu,\sigma^2)$:

$$\begin{aligned} EX &= \int_{-\infty}^{+\infty}\frac{x}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx \\ &= \int_{-\infty}^{+\infty}(\sigma y+\mu).\frac{1}{\sqrt{2\pi}}e^{-y^2/2}dy \\ &= \mu \end{aligned}$$

4. 均匀分布$X\sim U[a,b]$:

$$EX=\frac{a+b}{2}$$

5. 指数分布$X\sim Exp(\lambda)$:

$$EX=1/\lambda$$

例 3.1.2. 设$r.v.\ X$的分布律为

$$P\left(X=(-1)^k\frac{2^k}{k}\right)=\frac{1}{2^k},\quad k=1,2,\cdots$$

则X的数学期望不存在。

解: 由于

$$\sum_{k=1}^{\infty}|(-1)^k\frac{2^k}{k}|\frac{1}{2^k}=\sum_{k=1}^{\infty}\frac{1}{k}=+\infty$$

因此X的数学期望不存在。而尽管

$$\sum_{k=1}^{\infty}(-1)^k\frac{2^k}{k}\frac{1}{2^k}=\sum_{k=1}^{\infty}(-1)^k\frac{1}{k}=-ln2.$$

例 3.1.3. (Cauchy分布)设

$$p(x)=\frac{1}{\pi(1+x^2)},\quad x\in\mathcal{R},$$

则*: 该分布的期望不存在.*

解:容易看出,$p(x)$非负, 并且

$$\int_{-\infty}^{\infty}p(x)dx=\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{1}{1+x^2}dx=\frac{1}{\pi}\arctan x\,|_{-\infty}^{\infty}=1,$$

所以$p(x)$是一个密度函数(称为Cauchy分布), 但是

$$\int_{-\infty}^{\infty}|x|p(x)dx=\frac{2}{\pi}\int_{0}^{\infty}\frac{x}{1+x^2}dx=\infty,$$

所以Cauchy分布的期望不存在. #

§3.1.2 数学期望的性质

1. 若干个随机变量线性组合的期望, 等于各变量期望的线性组合. 假设$c_1,c_2,\ldots,c_n$为常数, 则有

$$E(c_1X_1+c_2X_2+\cdots+c_nX_n)=c_1EX_1+c_2EX_2+\cdots+c_nEX_n,$$

这里假定各变量的期望都存在.

例 3.1.4. 假设随机变量$X \sim B(n,p)$, 求EX.

解: 令$I_i \sim B(1,p)$, $i=1,2,\ldots,n$, 则$X=\sum_{i=1}^n I_i$ 且$EI_i=p$. 所以，$EX=\sum_{i=1}^n EI_i=np$.

2. 若干个独立随机变量之积的期望, 等于各变量的期望之积, 即

$$E(X_1X_2\cdots X_n)=EX_1EX_2\cdots EX_n,$$

这里假定各变量相互独立且期望都存在.

3. (随机变量函数的期望) 设随机变量X 为离散型, 有分布$P(X=a_i)=p_i$, $i=1,2,\ldots$, 或者为连续型, 有概率密度函数$f(x)$. 则

$$Eg(X)=\begin{cases} \sum_i g(a_i)p_i, & \sum_i |g(a_i)|p_i<\infty; \\ \int_{-\infty}^{+\infty} g(x)f(x)dx, & \int_{-\infty}^{+\infty} |g(x)|f(x)dx<\infty. \end{cases}$$

例 3.1.5. 假设c为常数, 则$EcX=cEX$.

例 3.1.6. 设随机变量$X \sim N(0,1)$, 求$Y=X^2+1$的数学期望.

解: 由$X \sim N(0,1)$, 我们有

$$\begin{aligned} EX^2 &= \int_{-\infty}^{+\infty} x^2.\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx \\ &= 1. \end{aligned}$$

所以, $EY=EX^2+1=2$.

例 3.1.7. 飞机场载客汽车上有*20*位乘客, 离开机场后共有*10*个车站可以下车, 若某个车站没有人下车则该车站不停车. 设乘客在每个车站下车的可能性相等, 以X表示停车的次数, 求EX.

解: 设

$$Y_i=\begin{cases} 1, & \text{第 } i \text{ 个车站有人下车} \\ 0, & \text{第 } i \text{ 个车站无人下车} \end{cases} \qquad i=1,\cdots,20.$$

则显然$X=\sum\limits_{i=1}^{20}Y_i$, 所以

$$\begin{aligned} EX &= \sum_{i=1}^{20} EY_i = \sum_{i=1}^{20} P(\text{第 } i \text{ 个车站有人下车}) \\ &= \sum_{i=1}^{20} [1-0.9^{20}] = 8.784. \end{aligned}$$

§3.1.3 条件期望

我们知道条件分布也是一个概率分布, 因此类似数学期望的定义, 我们可以给出条件期望的定义. 在给定了随机变量X取值x的条件之下, Y的条件期望, 我们记为$E(Y|X=x)$, 也可简记为$E(Y|x)$.

定义 3.1.3. *设X和Y为随机变量, 若(X,Y)为离散型, 且在给定$X=x$之下, Y 有分布$P(Y=a_i|X=x)=p_i$, $i=1,2,\ldots$, 或者(X,Y)为连续型, 且在给定$X=x$之下, Y的条件密度函数为$f(y|x)$. 则*

$$E(Y|X=x)=\begin{cases} \int_{-\infty}^{+\infty} yf(y|x)dy, & (X,Y) \text{ 为连续型;} \\ \sum_i a_i p_i, & (X,Y) \text{ 为离散型.} \end{cases}$$

期望所具有的性质条件期望同样满足.

例 3.1.8. *设$(X,Y)\sim N(a,b,\sigma_1^2,\sigma_2^2,\rho)$, 试计算$E(Y|X=x)$.*

解: 由于$Y|X=x\sim N(b+\rho\frac{\sigma_2}{\sigma_1}(x-a),(1-\rho^2)\sigma_2^2)$, 所以由二维正态分布的性质知$E(Y|X=x)=b+\rho\frac{\sigma_2}{\sigma_1}(x-a)$.

[注]: 条件期望$E(Y|X=x)$是x的函数, 当我们将x换为X时, $E(Y|X)$ 就是一个随机变量.

我们有如下的公式成立:

定理 3.1.1. *设X,Y为两个随机变量. 则有*

$$EX=E\{E[X|Y]\} \qquad [\text{全期望公式}]$$

证: 我们仅在连续型随机变量的情形下证明此定理. **证:** 我们仅在连续型随机变量的情形下证明此定理. 设X的p.d.f 为$f(x)$, Y的p.d.f 为$p(y)$, $X|Y=y$的p.d.f为$q(x|y)$. 则

$$\begin{aligned} EX &= \int_{-\infty}^{\infty} xf(x)dx = \int_{-\infty}^{\infty} x\int_{-\infty}^{\infty} q(x|y)p(y)dydx \\ &= \int\int_{-\infty}^{\infty} xq(x|y)dxp(y)dy = \int_{-\infty}^{\infty} E[X|Y=y]p(y)dy \\ &= E\{E[X|Y]\} \end{aligned}$$

[**推广**]: 当$g(X)$ 为可积随机变量时, 有$Eg(X)=E\{E[g(X)|Y]\}$.

由此得到求解期望的第二种方法: 先求解$h(x)=E(Y|X=x)$, 再求解$Eh(X)$, 即可求得EY.

例 3.1.9. *一窃贼被关在有3个门的地牢里, 其中第一个门通向自由. 出这门走3个小时便可以回到地面; 第2个门通向另一个地道, 走5个小时将返回到地牢; 第3个门通向更长的地道, 走7个小时也回到地牢. 若窃贼每次选择3个门的可能性总相同, 求他为获得自由而奔走的平均时间.*

解: 设这个窃贼需要走X小时才能到达地面, 并设Y代表他每次对3个门的选择情况, Y各以1/3的概率取值1, 2, 3. 则

$$EX = E[E(X|Y)] = \sum_{i=1}^{3} E(X|Y=i)P(Y=i)$$

注意到$E(X|Y=1)=3, E(X|Y=2)=5+EX, E(X|Y=3)=7+EX$, 所以

$$EX = \frac{1}{3}[3+5+EX+7+EX]$$

即得到$EX=15$.

例 3.1.10. *设$(X,Y)\sim N(a,b,\sigma_1^2,\sigma_2^2,\rho)$, 试计算$EXY$.*

解: 先算得

$$E(XY|X=x) = xE(Y|X=x) = x(b+\rho\frac{\sigma_2}{\sigma_1}(x-a));$$

所以

$$EXY = E(bX+\rho\frac{\sigma_2}{\sigma_1}X^2-\rho\frac{\sigma_2}{\sigma_1}aX)$$

$$
\begin{aligned}
&= ab + \rho\frac{\sigma_2}{\sigma_1}(a^2+\sigma_1^2) - \rho\frac{\sigma_2}{\sigma_1}a^2 \\
&= ab + \rho\sigma_1\sigma_2.
\end{aligned}
$$

§3.1.4 中位数

我们已经知道, 随机变量X的数学期望就是它的平均值, 因此从一定意义上, 数学期望刻画了随机变量所取之值的"中心位置". 但是, 我们也可以用别的数字特征来刻画随机变量的"中心位置". 中位数就是这样一种数字特征.

定义 3.1.4. 称μ为连续型随机变量X的中位数, 如果

$$P(X \le \mu) = \frac{1}{2}, \quad P(X \ge \mu) = \frac{1}{2}.$$

从定义上可以看出, m这个点把X 的分布从概率上一分两半: 在m左边占一半, m右边也占一半, 从概率上说, m 这个点正好居于中央, 这就是"中位数" 得名的由来. 在实用上, 中位数用得很多, 特别有不少社会统计资料, 常拿中位数来刻化某种量的代表性数值, 有时它比数学期望更说明问题, 例如, 某社区内人的收入的中位数告诉我们: 有一半人的收入低于此值, 另一半高于此值. 我们直观上感觉到这个值对该社区的收入情况, 的确很具有代表性, 和期望值相比它的一个优点是受个别特别大或特别小的值的影响很小, 而期望则不然, 举例而言, 若该社区中有一个收入在百万元以上, 则该社区的均值可能很高, 而绝大多数人并不富裕, 这个均值并不很有代表性, 中位数则不然, 它几乎不受少量这种特大值的影响.

从理论上说, 中位数与均值相比还与一个优点, 即它总存在, 而均值则不是对任何随机变量都存在. 虽则中位数有这些优点, 但在概率统计中, 无论理论和应用上, 数学期望的重要性都超过中位数, 其原因有一下两个方面:

1. 均值有很多优良的性质, 这些性质时使得在数学处理上很方便. 例如, $E(X_1+X_2) = EX_1 + EX_2$, 而X_1+X_2 的中位数与X_1, X_2各自的中位数之间, 不存在简单的联系, 这使中位数在数学上的处理很复杂且不方便;

2. 中位数本身固有的某些缺点：中位数可以不唯一, 且对于离散型随机变量不易定义.

例 3.1.11. 设随机变量$X \sim B(1, \frac{1}{2})$, 求X的中位数.

解: 由于X的分布函数为

$$F(x)=\begin{cases}0, & x\le 0\\ \frac{1}{2}, & 0<x<1\\ 1, & x\ge 1\end{cases}$$

由中位数的定义知区间(0,1)内的每一个数都是X的中位数, 所以此例说明中位数可以不唯一.

中位数的定义是p分位数定义的特例:

定义 3.1.5. 设$0<p<1$，称μ_p是随机变量ξ的p分位数，如果

$$P(\xi\le\mu_p)\ge p,\quad P(\xi\ge\mu_p)\ge 1-p.$$