Lecture 6

期望与方差

ruanyl@buaa.edu.cn
September 2016, Beihang University

随机变量的期望

离散随机变量及其函数的期望

离散随机变量可取至多可数个不同的实数, 因此它以写成有限或可数求和

$$X(\omega) = \sum_i x_i I_{A_i}(\omega) \tag{6.1}$$

这里 $A_i = \{\omega : X(\omega) = x_i\}$. $\bigcup_i A_i$ 形成概率空间的划分. $I_A(\omega)$ 是集合 A 上的指示函数,

$$I_A(\omega) = \begin{cases} 1, & \omega \in A \\ 0, & \omega \in A^c \end{cases}.$$

如果 X 仅取有限个值, 那么我们称 X 是**简单随机变量**.

定义 假设 X 是 (6.1) 中的离散随机变量. 定义如下“加权平均

值”

$$\sum_i x_i \cdot P(A_i) = x_1 P(A_1) + x_2 P(A_2) + x_3 P(A_3) + \cdots$$

如果这个求和是绝对收敛的, 那么我们把这个求和称为 X 的期望, 并且记作

$$E(X) = \sum_i x_i \cdot P(A_i) = \sum_i x_i \cdot P(X = x_i)$$

注意 $E(X)$ 的定义仅仅依赖于 X 的分布, 因此如果 X, Y 具有相同分布, 那么 $E(X) = E(Y)$.

定理 假设 $g(x)$ 是实数空间上某函数, X 是 (6.1) 中的离散随机变量. 那么 $Y = g(X)$ 的期望是 (如果该求和绝对收敛)

$$E(Y) = \sum_i g(x_i) \cdot P(X = x_i)$$

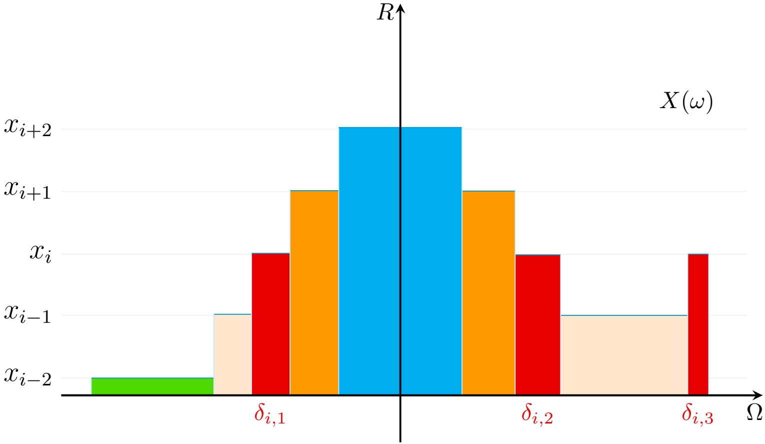
- 事实上

$$Y(\omega) = \sum_x g(x_i) \cdot I_{A_i}(\omega),$$

而 $\bigcup_i A_i = \bigcup_i \{X = x_i\}$ 形成概率空间的划分, 因此由定义可知 Y 的期望是

$$\sum_i g(x_i) \cdot P(A_i) = \sum_i g(x_i) \cdot P(X = x_i).$$

离散随机变量在概率空间的积分可以通过下图来理解



$$P\big(\omega : X(\omega) = x_i\big) = P\big(\delta_{i,1}\big) + P\big(\delta_{i,2}\big) + P\big(\delta_{i,3}\big)$$

Figure 6.1: 简单随机变量积分示意图

连续型随机变量及其函数的期望

我们已经知道如何求简单随机变量的积分. 为了求连续型随机变量的积分, 我们将用简单随机变量来近似连续型随机变量, 进而用简单随机变量的积分来近似连续型随机变量的积分.

给定随机变量 X, 将其取值空间 $\mathbb{R}$ 分割为互不相交的小区间 $B_i = [x_i, x_{i+1}), \bigcup_i B_i = \mathbb{R}$. 小区间的度量为

$$|B_i| = \Delta x_i = x_{i+1} - x_i.$$

如果划分充分细小, 我们可以认为随机变量 X 在 B_i 上近似地取常数值 $x_i \in B_i$. 例如, 如果 $X \geqslant 0$, 可取 $B_i = \left[\frac{i}{2^n}, \frac{i+1}{2^n}\right)$, 令

$$X_n(\omega) = \begin{cases} \frac{i}{2^n}, & X(\omega) \in \left[\frac{i}{2^n}, \frac{i+1}{2^n}\right), 0 \leqslant i \leqslant n2^n - 1 \\ n, & X(\omega) \geqslant n \end{cases}$$

那么 X_n 是一列收敛 $(n \to \infty)$ 到 X 的简单随机变量.

注意, 如果 $\bigcup_i B_i$ 是 $\mathbb{R}$ 的一个划分, 令 $\Omega_i = \{\omega : X(\omega) \in B_i\}$, 那么 $\bigcup_i \Omega_i$ 是 Ω 的一个划分.

一般随机变量在概率空间的积分可以通过下图来理解

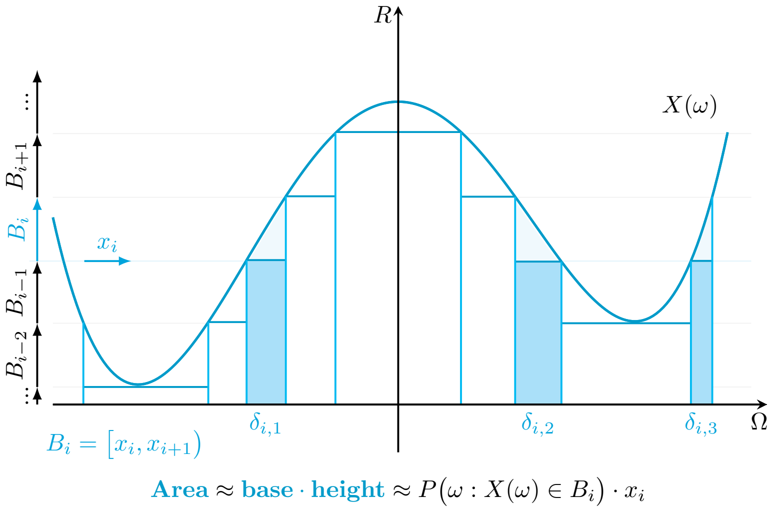


Figure 6.2: 随机变量积分示意图

因此我们只需要对简单随机变量 X_n 计算积分, 而 X_n 的积分(参见6.2)：

$$E\left(X_n\right)=\sum_i P\left(X_n=x_i\right)\cdot x_i=\sum_i P\left(X\in B_i\right)\cdot x_i$$

如果 X 的密度函数为 $f(x)$, 那么

$$\begin{aligned}
E\left(X_n\right)&=\sum_i P\left(X\in B_i\right)\cdot x_i\\
&=\sum_i \int_{-\infty}^{\infty} I_{B_i}\left(x\right)f\left(x\right)dx\cdot x_i\\
&=\int_{-\infty}^{\infty}\left(\sum_i x_i I_{B_i}\left(x\right)\right)f\left(x\right)dx\\
&=\int_{-\infty}^{\infty}\left(\sum_i x_i I_{B_i}\left(x\right)\right)f\left(x\right)dx
\end{aligned}$$

当 $n \to \infty$, 划分 $\bigcup_i B_i$ 变得越来越细,

$$\sum_i x_i I_{B_i}(x) \to x$$

并且

$$E(X_n) \to \int_{-\infty}^{\infty} x f(x)\, dx$$

因而我们有

定义 假设随机变量 X 的密度函数为 $f(x)$, 如果存在 (绝对收敛) 积分

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

那么把它称为 X 的期望, 并记作

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

定理 假设 $g(x)$ 是实数空间上某函数, X 是以 $f(x)$ 为密度函数的连续型随机变量. 那么 $Y=g(X)$ 的期望是 (如果该积分绝对收敛)

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

- 如同前面, 将 $g(x)$ 取值空间 $\mathbb{R}$ 分割为互不相交的小区间 $B_i=[y_i,y_{i+1})$, $\bigcup_i B_i=\mathbb{R}$. 小区间的度量为

$$|B_i|=\Delta y_i=y_{i+1}-y_i.$$

令 $C_i=g^{-1}(B_i)$. 于是 Y 可以由 $\bigcup_i C_i$ 上的简单随机变量 Y_n 近似 (参见6.3) 而 $E(Y)$ 可以用 $E(Y_n)$ 来近似

$$E(Y_n)=\sum_i P(Y_n=y_i)\cdot y_i$$

$$
\begin{aligned}
&= \sum_i P\left(Y \in B_i\right) \cdot y_i \\
&= \sum_i P\left(X \in g^{-1}\left(B_i\right)\right) \cdot y_i \\
&= \sum_i \int_{-\infty}^{\infty} I_{C_i}\left(x\right) f\left(x\right) dx \cdot y_i \\
&= \int_{-\infty}^{\infty} \left(\sum_i y_i I_{C_i}\left(x\right)\right) f\left(x\right) dx
\end{aligned}
$$

根据 Y_n 的定义,

$$
\sum_i y_i I_{C_i}\left(x\right) \to g\left(x\right)
$$

因而上面的求和收敛到

$$
E\left(Y_n\right) \to \int_{-\infty}^{\infty} g\left(x\right) f\left(x\right) dx
$$

对多个随机变量, 同样的结论成立, 例如已知随机向量 $\mathbf{Z} = (X, Y)$ 有分布密度 $f_{\mathbf{Z}}(\mathbf{z}) = f(x, y)$, $g(\mathbf{z}) = g(x, y)$ 是定义在 $\mathbb{R}^2$ 上的函数, 那么随机变量 $g(\mathbf{z})$ 的期望 (若存在) 为

$$E(\mathbf{Z}) = \begin{cases} \int\limits_{\mathbb{R}^2} g(\mathbf{z}) f_{\mathbf{Z}}(\mathbf{z})\, d\mathbf{z}, & \text{连续型} \\ \sum\limits_{\mathbf{z}} g(\mathbf{z}) f_{\mathbf{Z}}(\mathbf{z}), & \text{离散型} \end{cases}$$

$$= \begin{cases} \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} g(x, y) f(x, y)\, dxdy, & \text{连续型} \\ \sum\limits_{x} \sum\limits_{y} g(x, y) f(x, y), & \text{离散型} \end{cases}$$

随机变量的函数在概率空间的积分可以通过下图来理解

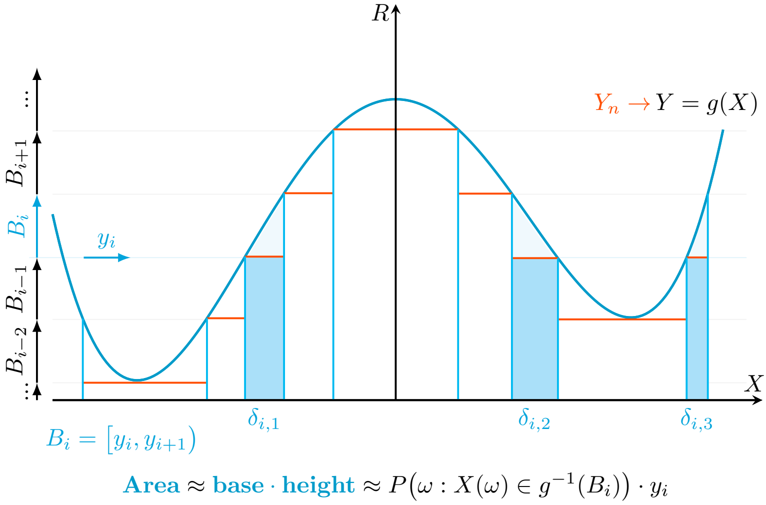


Figure 6.3: 随机变量函数的积分示意图

例题 已知 X 的密度为

$$f(x)=\begin{cases}\frac{3}{8}x^2, & 0<x<2\\ 0, & \text{其他}\end{cases}$$

试求 $\frac{1}{X^2}$ 的期望.

- $\frac{1}{X^2}$ 的期望为

$$E\left(\frac{1}{X^2}\right)=\int_{-\infty}^{\infty}\frac{1}{x^2}f(x)\,dx=\int_0^2\frac{1}{x^2}\cdot\frac{3}{8}x^2dx=\frac{3}{4}$$

条件期望

在条件概率一节我们看到给定 $X=x, f_Y(\cdot|x)$ 是一个概率密度, $g(x)$ 是实数空间上某函数, 因此可以考虑随机变量 $g(Y)$ 在

给定 $X=x$ 时的期望, 称之为**条件期望**, 写作

$$E\left(g\left(Y\right)|X=x\right)=\begin{cases}\int\limits_{-\infty}^{\infty}g\left(y\right)f_Y\left(y|x\right)dy, & \text{连续型}\\ \sum\limits_{y}g\left(y\right)f_Y\left(y|x\right), & \text{离散型}\end{cases}$$

此时它是一个 x 的函数, 对上式同乘以 $f_X\left(x\right)$ 并积分 (求和) 便得到

$$E\left(g\left(Y\right)\right)=\begin{cases}\int\limits_{-\infty}^{\infty}E\left(g\left(Y\right)|X=x\right)f_X\left(x\right)dx, & \text{连续型}\\ \sum\limits_{x}E\left(g\left(Y\right)|X=x\right)f_X\left(x\right), & \text{离散型}\end{cases} \tag{6.2}$$

这是条件期望的性质, 常常用于期望的计算.

待续