第四章 随机变量的深入讨论

Probability and Statistics for Computer Scientists

第四章 随机变量的深入讨论

- 随机变量函数的分布密度函数
- 协方差和相关
- 再论条件期望和条件方差
- 矩母函数
- 随机数个相互独立随机变量之和

§1 随机变量函数的分布密度函数

导出的密度函数

连续随机变量 X 的函数 $Y = g(X)$ 的分布密度函数

(1) 使用如下公式计算 Y 的分布函数 (CDF) F_Y

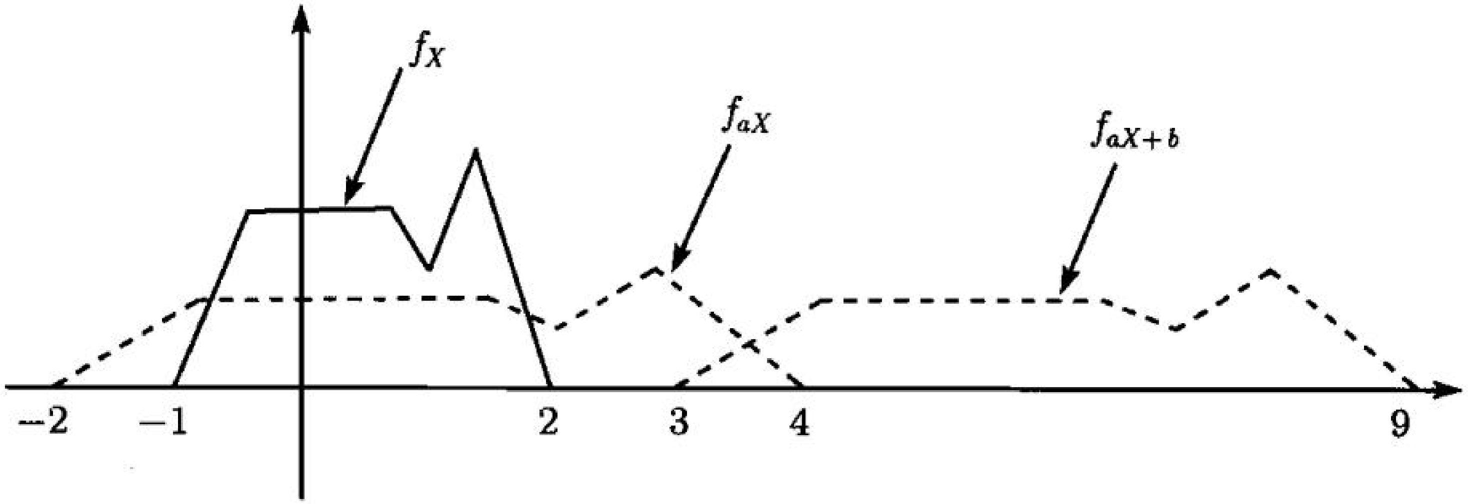
$$F_Y(y) = \mathrm{P}(g(X) \leqslant y) = \int_{\{x|g(x)\leqslant y\}} f_X(x)\mathrm{d}x.$$

(2) 对 F_Y 求导, 得到 Y 的 PDF:

$$f_Y(y) = \frac{\mathrm{d}F_Y}{\mathrm{d}y}(y).$$

§1 随机变量函数的分布密度函数

1.1 Y是X的线性函数



图中 $a=2,\, b=5.$ $\quad f_Y(y)=\dfrac{1}{|a|}f_X\left(\dfrac{y-b}{a}\right).$

§1 随机变量函数的分布密度函数

证明该公式

$$
\begin{aligned}
F_Y(y) &= \mathrm{P}(Y \leqslant y) \\
&= \mathrm{P}(aX + b \leqslant y) \\
&= \mathrm{P}\left(X \leqslant \frac{y-b}{a}\right) \\
&= F_X\left(\frac{y-b}{a}\right)
\end{aligned}
$$

对上述等式微分, 运用复合函数求导方法, 可得

$$
f_Y(y) = \frac{\mathrm{d}F_Y}{\mathrm{d}y}(y) = \frac{1}{a} f_X\left(\frac{y-b}{a}\right)
$$

§1 随机变量函数的分布密度函数

例 4.4 (指数随机变量的线性函数) 假设随机变量 X 服从参数为 λ 的指数分布，密度函数为

$$f_X(x)=\begin{cases}\lambda e^{-\lambda x}, & 若\ x\geqslant 0,\\ 0, & 其他,\end{cases}$$

其中 λ 是正的参数. 定义 $Y=aX+b$, 则

$$f_Y(y)=\begin{cases}\dfrac{\lambda}{|a|}e^{-\lambda(y-b)/a}, & 若\ (y-b)/a\geqslant 0,\\ 0, & 其他.\end{cases}$$

§1 随机变量函数的分布密度函数

1.2 Y是X的单调函数

(a) 严格单调递增：对任意的 $x, x' \in I$, 满足 $x < x'$, 则 $g(x) < g(x')$;
(b) 严格单调递减：对任意的 $x, x' \in I$, 满足 $x < x'$, 则 $g(x) > g(x')$.

假设 g 是严格单调函数, 其逆函数 h 满足: 对 X 的取值空间内任意一点 x,

$$y = g(x) \quad \text{当且仅当} \quad x = h(y).$$

而且函数 h 是可微的, 则 Y 在支撑集 $\{y : f_Y(y) > 0\}$ 内的密度函数是

$$f_Y(y) = f_X(h(y)) \left| \frac{\mathrm{d}h}{\mathrm{d}y}(y) \right|.$$

§1 随机变量函数的分布密度函数

假设 g 是严格递增函数. 则

$$F_Y(y)=\mathrm{P}(g(X)\leqslant y)=\mathrm{P}(X\leqslant h(y))=F_X(h(y)),$$

$$f_Y(y)=\frac{\mathrm{d}F_Y}{\mathrm{d}y}(y)=f_X(h(y))\frac{\mathrm{d}h}{\mathrm{d}y}(y).$$

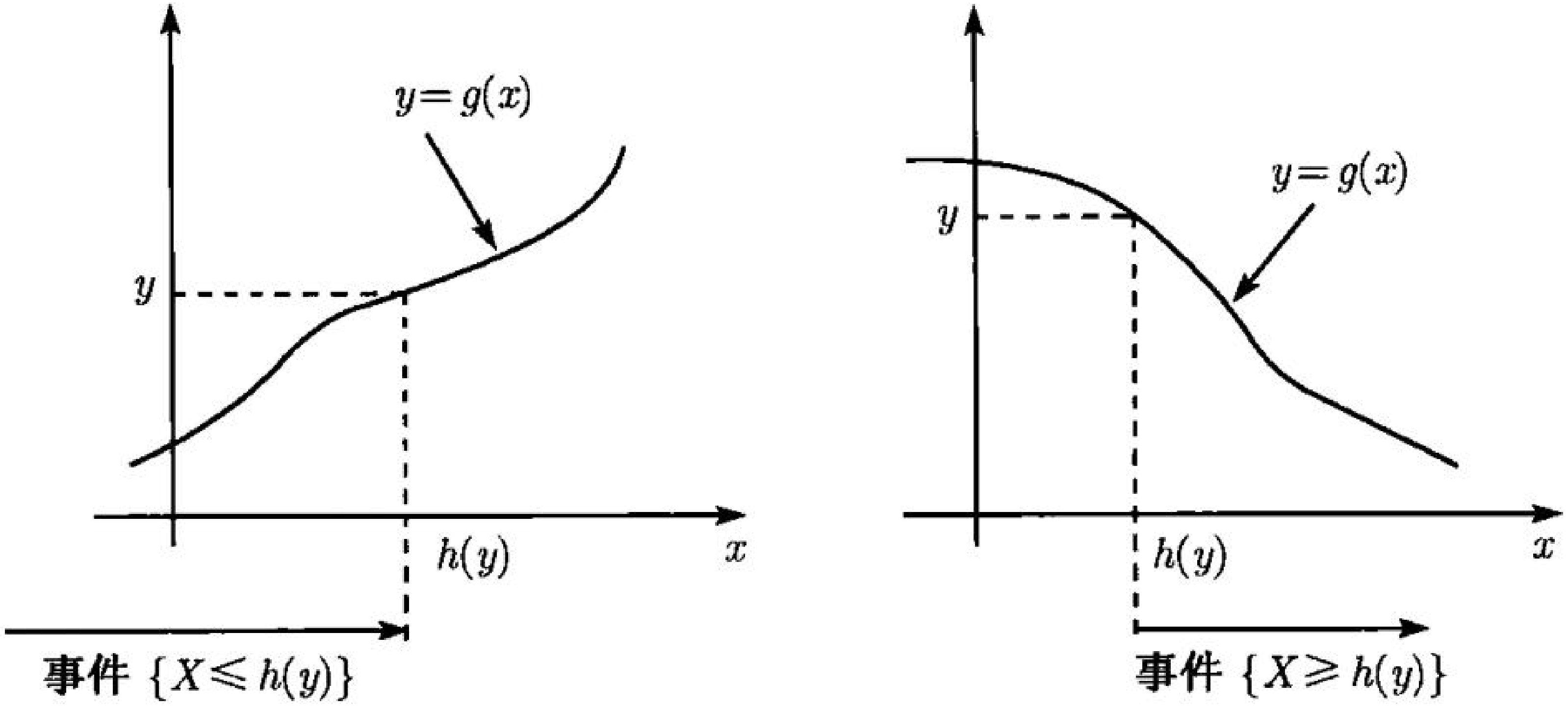
因为 g 是严格递增时, 函数 h 也是严格递增的, 所以它的导数是非负的:

$$\frac{\mathrm{d}h}{\mathrm{d}y}(y)=\left|\frac{\mathrm{d}h}{\mathrm{d}y}(y)\right|$$

当 g 是单调递减时, $F_Y(y)=\mathrm{P}(g(X)\leqslant y)=\mathrm{P}(X\geqslant h(y))=1-F_X(h(y))$,

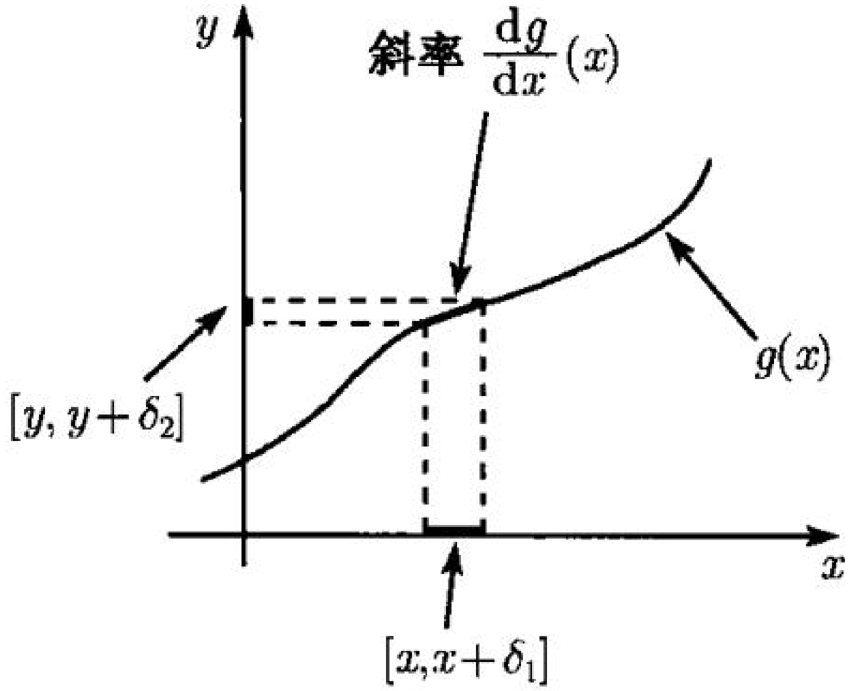
推导过程类似

§1 随机变量函数的分布密度函数



§1 随机变量函数的分布密度函数

概率密度函数，即随机变量落入小区间的概率



g 在点 x 处的斜率 $\frac{\delta_2}{\delta_1} \approx \frac{dg}{dx}(x)$

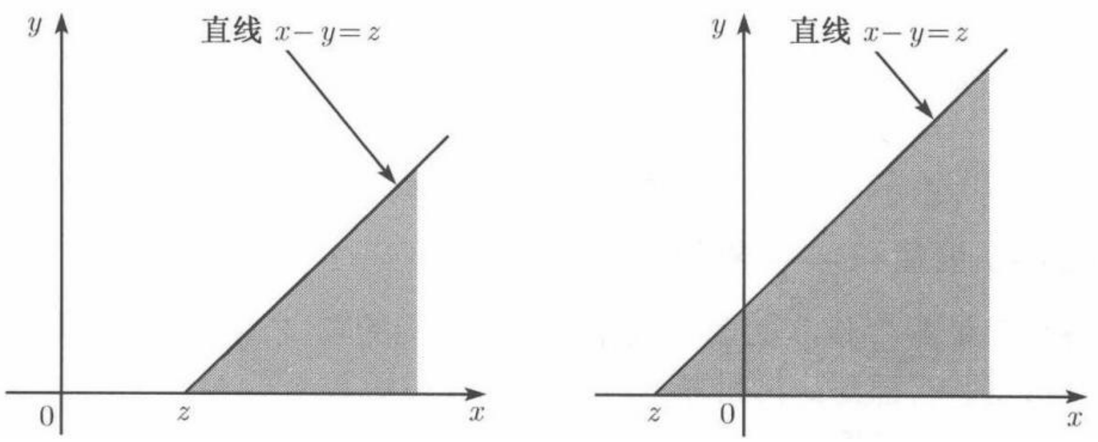
逆函数 $\frac{\delta_1}{\delta_2} \approx \frac{dh}{dy}(y)$

$$f_Y(y)\delta_2 \approx P(y \leqslant Y \leqslant y+\delta_2) = P(x \leqslant X \leqslant x+\delta_1) \approx f_X(x)\delta_1.$$

§1 随机变量函数的分布密度函数

1.3 两个随机变量的函数

例 4.9 罗密欧和朱丽叶定期约会，他们每个人每次到达约会地点时都会离约定的时间有延迟，而且他们的延迟时间是彼此相互独立的。假定延迟的时间都服从指数分布，参数为 λ。那么他们到达约会地点的时间差具有什么样的概率密度函数？



§1 随机变量函数的分布密度函数

$Z = X - Y$

当 $z \geqslant 0$ $F_Z(z) = \mathrm{P}(X - Y \leqslant z) = 1 - \mathrm{P}(X - Y > z)$

$$
\begin{aligned}
&= 1 - \int_0^{\infty} \mathrm{d}y \int_{z+y}^{\infty} f_{X,Y}(x, y) \mathrm{d}x \\
&= 1 - \int_0^{\infty} \lambda \mathrm{e}^{-\lambda y} \mathrm{d}y \int_{z+y}^{\infty} \lambda \mathrm{e}^{-\lambda x} \mathrm{d}x \\
&= 1 - \int_0^{\infty} \lambda \mathrm{e}^{-\lambda y} \mathrm{e}^{-\lambda(z+y)} \mathrm{d}y \\
&= 1 - \mathrm{e}^{-\lambda z} \int_0^{\infty} \lambda \mathrm{e}^{-2\lambda y} \mathrm{d}y \\
&= 1 - \frac{1}{2} \mathrm{e}^{-\lambda z}
\end{aligned}
$$

§1 随机变量函数的分布密度函数

当 $z < 0$

$$F_Z(z) = \mathrm{P}(Z \leqslant z) = \mathrm{P}(-Z \geqslant -z) = \mathrm{P}(Z \geqslant -z) = 1 - F_Z(-z).$$

$-z > 0$

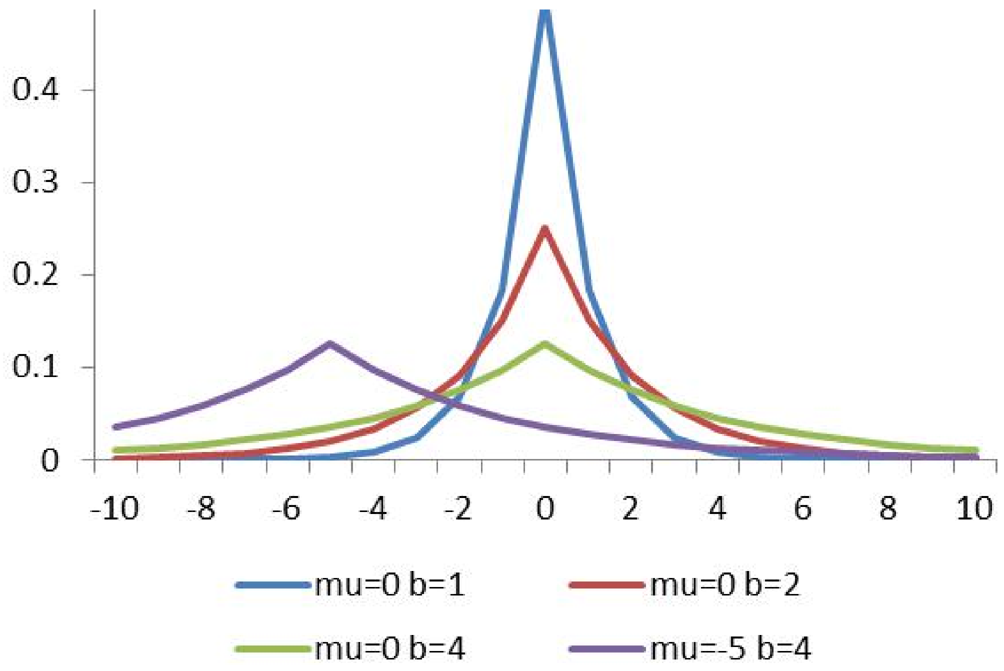
$$F_z(z) = 1 - F_Z(-z) = 1 - \left(1 - \frac{1}{2}\mathrm{e}^{-\lambda(-z)}\right) = \frac{1}{2}\mathrm{e}^{\lambda z}.$$

对分布函数进行微分, 可以得到密度函数, 即

$$f_Z(z) = \begin{cases} \dfrac{\lambda}{2}\mathrm{e}^{-\lambda z}, & \text{若 } z \geqslant 0 \text{ 时}, \\ \dfrac{\lambda}{2}\mathrm{e}^{\lambda z}, & \text{若 } z < 0 \text{ 时}. \end{cases}$$

双边指数密度函数

§1 随机变量函数的分布密度函数

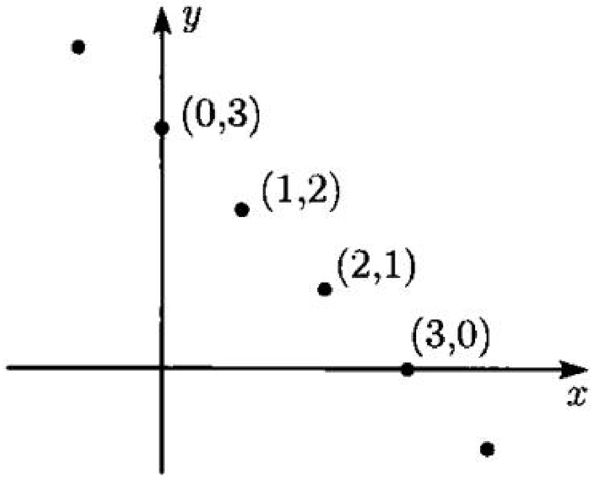


§1 随机变量函数的分布密度函数

1.4 独立随机变量和-----卷积

X 和 Y 都是离散的情况下, $Z=X+Y$ 的分布.

$$\begin{aligned} p_Z(z) &= \mathrm{P}(X+Y=z) \\ &= \sum_{\{(x,y)|x+y=z\}} \mathrm{P}(X=x, Y=y) \\ &= \sum_x \mathrm{P}(X=x, Y=z-x) \\ &= \sum_x p_X(x) p_Y(z-x). \end{aligned}$$



§1 随机变量函数的分布密度函数

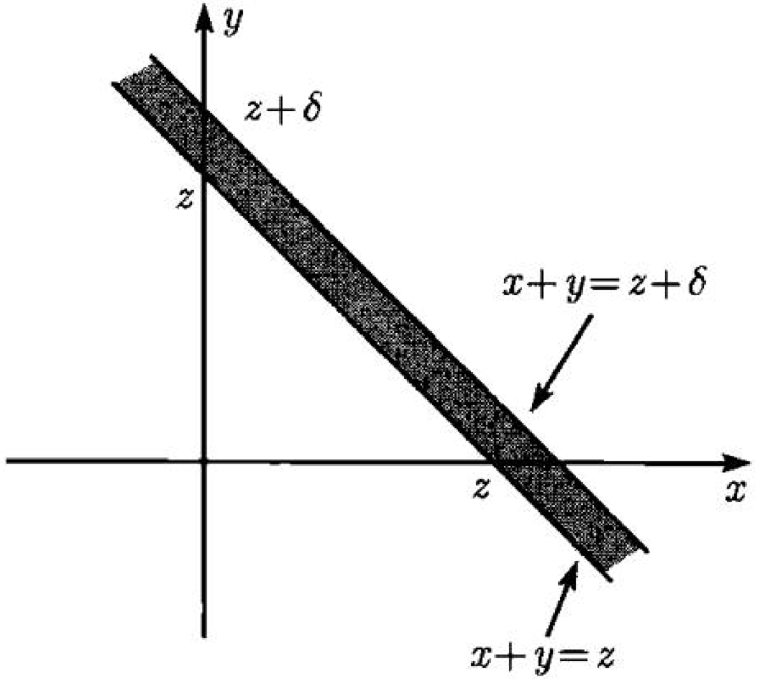
X 和 Y 为独立的连续型随机变量，

$$\begin{aligned}
\mathrm{P}(Z \leqslant z|X=x) &= \mathrm{P}(X+Y \leqslant z|X=x) \\
&= \mathrm{P}(x+Y \leqslant z) \\
&= \mathrm{P}(Y \leqslant z-x),
\end{aligned}$$

$$f_{X,Z}(x,z) = f_X(x) f_{Z|X}(z|x) = f_X(x) f_Y(z-x),$$

$$f_Z(z) = \int_{-\infty}^{\infty} f_{X,Z}(x,z)\mathrm{d}x = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)\mathrm{d}x.$$

§1 随机变量函数的分布密度函数



$$\mathrm{P}(z \leqslant X+Y \leqslant z+\delta) \approx f_Z(z)\delta$$

$$\begin{aligned} f_Z(z)\delta &= \mathrm{P}(z \leqslant X+Y \leqslant z+\delta) \\ &= \int_{-\infty}^{\infty}\int_{z-x}^{z-x+\delta} f_X(x)f_Y(y)\mathrm{d}y\mathrm{d}x \\ &\approx \int_{-\infty}^{\infty} f_X(x)f_Y(z-x)\delta\mathrm{d}x. \end{aligned}$$

§1 随机变量函数的分布密度函数

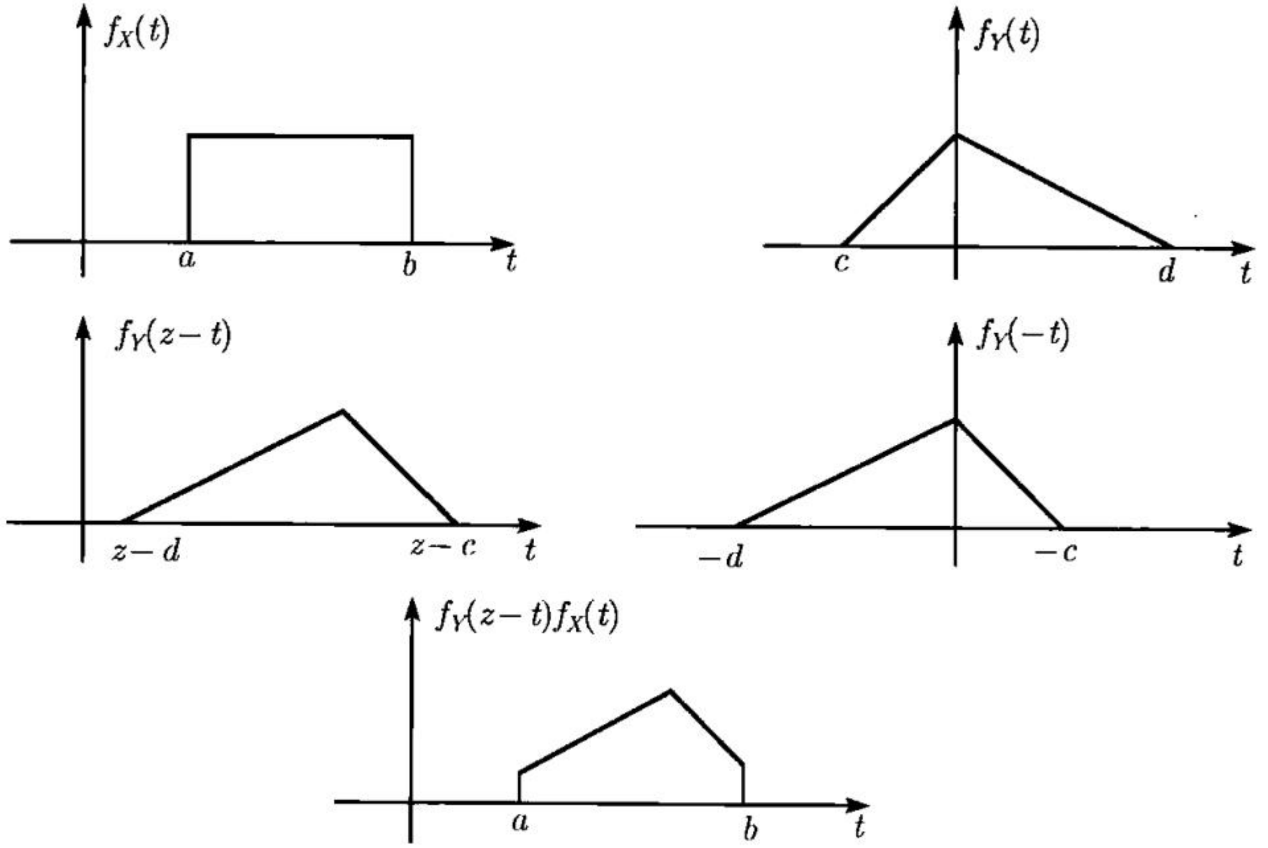
1.5 卷积的图像计算法

考虑概率密度函数 $f_X(t)$ 和 $f_Y(t)$ 给定 z 一个值, 计算卷积

$$f_Z(z)=\int_{-\infty}^{\infty} f_X(t)f_Y(z-t)\mathrm{d}t$$

§1 随机变量函数的分布密度函数

卷积是两个变量在某范围内相乘后求和的结果。



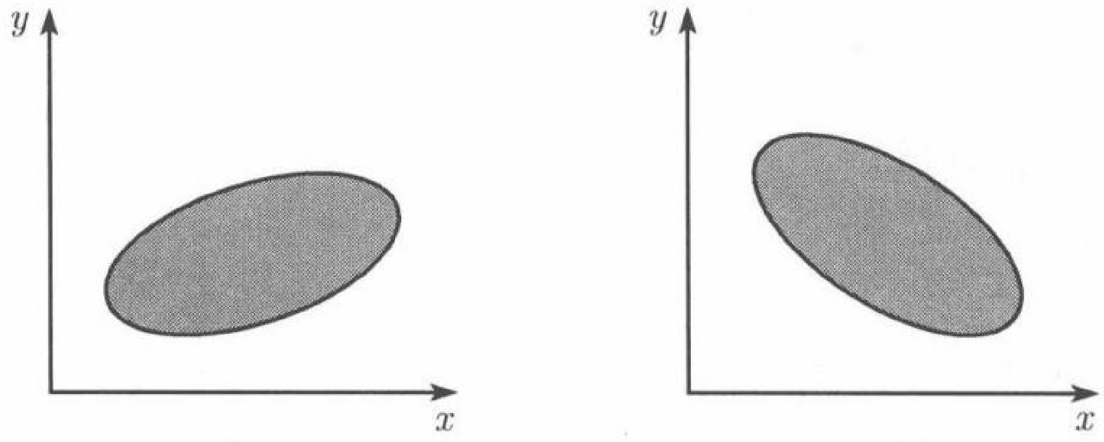
§2 协方差和相关

2.1 X和Y的协方差

$$\operatorname{cov}(X, Y) = \mathrm{E}[(X - \mathrm{E}[X])(Y - \mathrm{E}[Y])].$$

$$\operatorname{cov}(X, Y) = \mathrm{E}[XY] - \mathrm{E}[X]\mathrm{E}[Y]$$

当 $\operatorname{cov}(X, Y) = 0$ 时, 我们说 X 和 Y 是不相关的.



(a) (b)

§2 协方差和相关

协方差COV(X,Y)的性质

$$\operatorname{cov}(X, X) = \operatorname{var}(X),$$

$$\operatorname{cov}(X, aY + b) = a \cdot \operatorname{cov}(X, Y)$$

$$\operatorname{cov}(X, Y + Z) = \operatorname{cov}(X, Y) + \operatorname{cov}(X, Z).$$

§2 协方差和相关

两个方差非零的随机变量 X 和 Y 的相关系数

$$\rho(X,Y)=\frac{\mathrm{cov}(X,Y)}{\sqrt{\mathrm{var}(X)\mathrm{var}(Y)}}$$

可以证明 $\rho=1$ $(\rho=-1)$ 当且仅当存在一个正的 (负的) 常数 c, 使得

$$Y-\mathrm{E}[Y]=c(X-\mathrm{E}[X])$$

§2 协方差和相关

例 4.14 考虑一个硬币的 n 次独立的抛掷，其中正面朝上的概率是 p. 设 X 和 Y 分别是正面朝上和负面朝上的次数，现在让我们来看一下 X 和 Y 的相关系数.

我们总有 $X+Y=n$ 且 $\mathrm{E}[X]+\mathrm{E}[Y]=n$. 因此

$$X-\mathrm{E}[X]=-(Y-\mathrm{E}[Y]).$$

$$\begin{aligned}\operatorname{cov}(X,Y)&=\mathrm{E}\Big[(X-\mathrm{E}[X])(Y-\mathrm{E}[Y])\Big]\\&=-\mathrm{E}\Big[(X-\mathrm{E}[X])^2\Big]\\&=-\operatorname{var}(X).\end{aligned}$$

$$\rho(X,Y)=\frac{\operatorname{cov}(X,Y)}{\sqrt{\operatorname{var}(X)\operatorname{var}(Y)}}=\frac{-\operatorname{var}(X)}{\sqrt{\operatorname{var}(X)\operatorname{var}(X)}}=-1.$$

§2 协方差和相关

2.2 随机变量和的方差

设随机变量 $X_1, \cdots, X_n$ 具有有限的方差, 则

$$\mathrm{var}(X_1 + X_2) = \mathrm{var}(X_1) + \mathrm{var}(X_2) + 2\mathrm{cov}(X_1, X_2),$$

$$\mathrm{var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{var}(X_i) + \sum_{\{(i,j)|i\neq j\}} \mathrm{cov}(X_i, X_j).$$

§2 协方差和相关

上述公式, 可以如下推导: 简记 $\tilde{X}_i = X_i - \mathrm{E}[X_i]$,

$$\begin{aligned}
\mathrm{var}\left(\sum_{i=1}^{n}\tilde{X}_i\right) &= \mathrm{E}\left[\left(\sum_{i=1}^{n}\tilde{X}_i\right)^2\right] \\
&= \mathrm{E}\left[\sum_{i=1}^{n}\sum_{j=1}^{n}\tilde{X}_i\tilde{X}_j\right] \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n}\mathrm{E}[\tilde{X}_i\tilde{X}_j] \\
&= \sum_{i=1}^{n}\mathrm{E}[\tilde{X}_i^2] + \sum_{\{(i,j)|i\neq j\}}\mathrm{E}[\tilde{X}_i\tilde{X}_j] \\
&= \sum_{i=1}^{n}\mathrm{var}(X_i) + \sum_{\{(i,j)|i\neq j\}}\mathrm{cov}(X_i, X_j).
\end{aligned}$$

§2 协方差和相关

例 4.15 考虑 2.5 节中讨论的帽子问题. 有 n 个人将帽子扔进一个盒子, 然后每人随机地选一顶帽子. 设 X 是拿到自己帽子的人数, 现在计算 X 的方差. 设 X_i 表

$X_i = 1$, 表示拿到了自己的帽子.

$$X = X_1 + \cdots + X_n.$$

$$\mathrm{E}[X_i] = \frac{1}{n}, \quad \mathrm{var}(X_i) = \frac{1}{n}\left(1 - \frac{1}{n}\right).$$

$$\begin{aligned}
\mathrm{cov}(X_i, X_j) &= \mathrm{E}[X_i X_j] - \mathrm{E}[X_i]\mathrm{E}[X_j] \\
&= \mathrm{P}(X_i = 1, X_j = 1) - \frac{1}{n} \cdot \frac{1}{n} \\
&= \mathrm{P}(X_i = 1)\mathrm{P}(X_j = 1 | X_i = 1) - \frac{1}{n^2} \\
&= \frac{1}{n} \cdot \frac{1}{n-1} - \frac{1}{n^2} \\
&= \frac{1}{n^2(n-1)}.
\end{aligned}$$

$$\begin{aligned}
\mathrm{var}(X) &= \mathrm{var}\left(\sum_{i=1}^{n} X_i\right) \\
&= \sum_{i=1}^{n} \mathrm{var}(X_i) + \sum_{\{(i,j) | i \neq j\}} \mathrm{cov}(X_i, X_j) \\
&= n \cdot \frac{1}{n}\left(1 - \frac{1}{n}\right) + n(n-1) \cdot \frac{1}{n^2(n-1)} \\
&= 1.
\end{aligned}$$

§3 再论条件期望和条件方差

例 4.16 假设我们在投掷一个不均匀的硬币，正面朝上的概率，记为 Y，也是随机的. 假定正面朝上的概率 Y 的分布为已知，它是 $[0,1]$ 上的分布. 现在我们投掷 n 次硬币，定义 X 为正面朝上的总次数. 由于对任意的 $y \in [0,1]$，我们有 $\mathrm{E}[X|Y=y]=ny$，所以 $\mathrm{E}[X|Y]$ 是随机变量 nY.

$$\mathrm{E}[\mathrm{E}[X|Y]] = \begin{cases} \sum_y \mathrm{E}[X|Y=y]p_Y(y), & Y \text{ 离散}, \\ \int_{-\infty}^{\infty} \mathrm{E}[X|Y=y]f_Y(y)\mathrm{d}y, & Y \text{ 连续}. \end{cases}$$

重期望法则：$\mathrm{E}[\mathrm{E}[X|Y]] = \mathrm{E}[X]$.

§3 再论条件期望和条件方差

重期望法则应用举例

例 4.18 (全班平均成绩与分组平均) 一个班级有 n 名学生. 学生 i 的测验分数记为 x_i. 已知班级测验的平均分为

$$m = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

现将全部学生分成 k 个互不相交的子集 $A_1, \cdots, A_k$(组).

$$m_s = \frac{1}{n_s}\sum_{i\in A_s} x_i.$$

§3 再论条件期望和条件方差

全班的平均分数可以用每组的平均分数 m_s 的加权平均来计算,

$$
\begin{aligned}
\sum_{s=1}^{k} \frac{n_s}{n} m_s &= \sum_{s=1}^{k} \frac{n_s}{n} \cdot \frac{1}{n_s} \sum_{i \in A_s} x_i \\
&= \frac{1}{n} \sum_{s=1}^{k} \sum_{i \in A_s} x_i \\
&= \frac{1}{n} \sum_{i=1}^{n} x_i \\
&= m.
\end{aligned}
$$

§3 再论条件期望和条件方差

用重期望法则来考虑

随机地选择一位学生,

$X=$ 被选中的学生的成绩;

$Y=$ 被选中的学生所在的组, $(Y\in\{1,\cdots,k\})$.

$$\mathrm{E}[X]=m$$

$$\mathrm{E}[X|Y=s]=\frac{1}{n_s}\sum_{i\in A_s}x_i=m_s.$$

$$m=\mathrm{E}[X]=\mathrm{E}[\mathrm{E}[X|Y]]=\sum_{s=1}^{k}\mathrm{E}[X|Y=s]\mathrm{P}(Y=s)=\sum_{s=1}^{k}\frac{n_s}{n}m_s.$$

§3 再论条件期望和条件方差

3.1 条件期望作为估计量

将 Y 视为能提供 X 信息的观测值,将条件期望作为给定 Y 的条件下,对 X 的估计,

$$\hat{X} = \mathrm{E}[X|Y].$$

估计误差 $\tilde{X} = \hat{X} - X.$

估计误差没有系统性的正或负的偏倚.

$$\mathrm{E}[\tilde{X}] = \mathrm{E}[\tilde{X}|Y] = \mathrm{E}[(\hat{X} - X)|Y] = \mathrm{E}[\hat{X}|Y] - \mathrm{E}[X|Y] = \hat{X} - \hat{X} = 0.$$

$\hat{X}$与估计误差 $\tilde{X}$ 是不相关的

$$\mathrm{E}[\hat{X}\tilde{X}] = \mathrm{E}[\mathrm{E}[\hat{X}\tilde{X}|Y]] = \mathrm{E}[\hat{X}\mathrm{E}[\tilde{X}|Y]] = 0, \quad \mathrm{E}[\hat{X}\tilde{X}|Y] = \hat{X}\mathrm{E}[\tilde{X}|Y] = 0.$$

$$\mathrm{cov}(\hat{X}, \tilde{X}) = \mathrm{E}[\hat{X}\tilde{X}] - \mathrm{E}[\hat{X}]\mathrm{E}[\tilde{X}] = 0 - \mathrm{E}[X] \cdot 0 = 0.$$

§3 再论条件期望和条件方差

3.2 条件方差

$$\mathrm{var}(X|Y) = \mathrm{E}[(X - \mathrm{E}[X|Y])^2|Y] = \mathrm{E}[\tilde{X}^2|Y].$$

在已知 $\{Y = y\}$ 的条件下, X的条件方差为

$$\mathrm{var}(X|Y = y) = \mathrm{E}[\tilde{X}^2|Y = y].$$

估计误差

$$\mathrm{var}(\tilde{X}) = \mathrm{E}[\tilde{X}^2] = \mathrm{E}[\mathrm{E}[\tilde{X}^2|Y]] = \mathrm{E}[\mathrm{var}(X|Y)],$$

全方差法则：$\mathbf{var}(X) = \mathrm{E}[\mathbf{var}(X|Y)] + \mathbf{var}(\mathrm{E}[X|Y]).$

§3 再论条件期望和条件方差

例 4.20 (学生成绩的方差与分组方差) 所讨论的问题背景与例 4.18 中的相同, 我们重新考虑这些随机变量

$$X = \text{学生的成绩},$$

$$Y = \text{该生所在的组}, \qquad (Y \in \{1, \cdots, k\}).$$

对于 $\operatorname{var}(X) = \mathrm{E}[\operatorname{var}(X|Y)] + \operatorname{var}(\mathrm{E}[X|Y])$ 每组内部方差的平均数+各组之间的方差

$$\mathrm{E}[\operatorname{var}(X|Y)] = \sum_{s=1}^{k} \mathrm{P}(Y = s)\operatorname{var}(X|Y = s) = \sum_{s=1}^{k} \frac{n_s}{n}\operatorname{var}(X|Y = s),$$

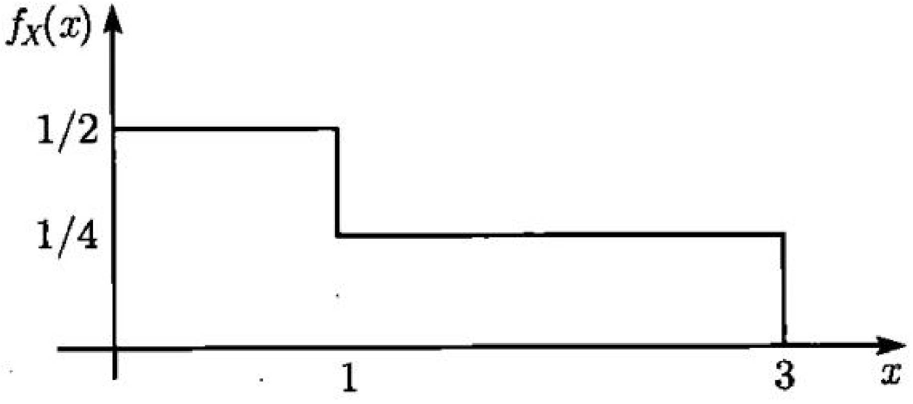
$\operatorname{var}(\mathrm{E}[X|Y])$ 就是各组均值波动性的度量.

§3 再论条件期望和条件方差

例 4.21 (通过给定条件来计算方差) 考虑一个连续随机变量 X, 它的概率密度函数在图 4.13 中给出, 我们定义一个辅助的随机变量 Y 如下:

$$Y = \begin{cases} 1, & 若\ x < 1, \\ 2, & 若\ x \geqslant 1. \end{cases}$$

这里, $\mathrm{E}[X|Y]$ 以 1/2 的概率分别取值 2 和 1/2. 因此, $\mathrm{E}[X|Y]$ 的均值为 5/4.



§3 再论条件期望和条件方差

$$\mathrm{var}(\mathrm{E}[X|Y]) = \frac{1}{2}\left(\frac{1}{2}-\frac{5}{4}\right)^2 + \frac{1}{2}\left(2-\frac{5}{4}\right)^2 = \frac{9}{16}.$$

当在给定 $Y=1$ 或 $Y=2$ 的条件下, X 在长度为 1 或 2 的线段上均匀分布.

因此

$$\mathrm{var}(X|Y=1) = \frac{1}{12}, \qquad \mathrm{var}(X|Y=2) = \frac{4}{12},$$

且

$$\mathrm{E}[\mathrm{var}(X|Y)] = \frac{1}{2}\cdot\frac{1}{12} + \frac{1}{2}\cdot\frac{4}{12} = \frac{5}{24}.$$

归总, 得

$$\mathrm{var}(X) = \mathrm{E}[\mathrm{var}(X|Y)] + \mathrm{var}(\mathrm{E}[X|Y]) = \frac{5}{24} + \frac{9}{16} = \frac{37}{48}.$$

§4 矩母函数

4.1 矩母函数

$$M_X(s) = \mathrm{E}[\mathrm{e}^{sX}].\quad \text{简记为 } M(s).$$

$$\begin{cases} M(s) = \sum_x \mathrm{e}^{sx} p_X(x), \\ M(s) = \int_{-\infty}^{\infty} \mathrm{e}^{sx} f_X(x)\mathrm{d}x. \end{cases}$$

§4 矩母函数

对于矩母函数：通过这个函数，可以求出所有n阶矩。

$$e^{sx} = 1 + sx + \frac{s^2x^2}{2!} + \frac{s^3x^3}{3!} + \cdots + \frac{s^nx^n}{n!}$$

$$M_X(s) = E[e^{sX}] = 1 + sE[X] + \frac{s^2E[X^2]}{2!} + \frac{s^3E[X^3]}{3!} + \cdots + \frac{s^nE[X^n]}{n!}$$

$$= 1 + sm_1 + \frac{s^2m_2}{2!} + \frac{s^3m_3}{3!} + \cdots + \frac{s^nm_n}{n!}$$

2阶矩　3阶矩　n阶矩.

§4 矩母函数

例 4.23 (泊松随机变量的矩母函数) 设随机变量 X 服从参数为 λ 的泊松分布:

$$p_X(x) = \frac{\lambda^x e^{-\lambda}}{x!}, \quad x = 0, 1, 2, \cdots,$$

则其矩母函数如下所示

$$M(s) = \sum_{x=0}^{\infty} e^{sx} \frac{\lambda^x e^{-\lambda}}{x!}.$$

记 $a = e^s \lambda$, 则

$$M(s) = e^{-\lambda} \sum_{x=0}^{\infty} \frac{a^x}{x!} = e^{-\lambda} e^a = e^{a-\lambda} = e^{\lambda(e^s - 1)}.$$

§4 矩母函数

例 4.24 (指数随机变量的矩母函数) 设随机变量 X 服从参数为 λ 的指数分布

$$f_X(x)=\lambda \mathrm{e}^{-\lambda x}, \qquad x\geqslant 0,$$

则

$$\begin{aligned} M(s) &= \lambda\int_0^{\infty} \mathrm{e}^{sx}\mathrm{e}^{-\lambda x}\mathrm{d}x \\ &= \lambda\int_0^{\infty} \mathrm{e}^{(s-\lambda)x}\mathrm{d}x \\ &= \lambda\frac{\mathrm{e}^{(s-\lambda)x}}{s-\lambda}\bigg|_0^{\infty} \qquad (\text{当 } s<\lambda \text{ 时}) \\ &= \frac{\lambda}{\lambda-s}. \end{aligned}$$

§ 4 矩母函数

4.2 从矩母函数到矩

$$M(s)=\int_{-\infty}^{\infty}\mathrm{e}^{sx}f_X(x)\mathrm{d}x,$$

在 $M(s)$ 定义式两边取 s 的导数

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}s}M(s)&=\frac{\mathrm{d}}{\mathrm{d}s}\int_{-\infty}^{\infty}\mathrm{e}^{sx}f_X(x)\mathrm{d}x\\&=\int_{-\infty}^{\infty}\frac{\mathrm{d}}{\mathrm{d}s}\mathrm{e}^{sx}f_X(x)\mathrm{d}x\\&=\int_{-\infty}^{\infty}x\mathrm{e}^{sx}f_X(x)\mathrm{d}x.\end{aligned}$$

上面的等式对 s 取任何值都成立①. 考虑 $s=0$ 时的特殊情况, 有

$$\frac{\mathrm{d}}{\mathrm{d}s}M(s)\Big|_{s=0}=\int_{-\infty}^{\infty}xf_X(x)\mathrm{d}x=\mathrm{E}[X].$$

§4 矩母函数

更广泛地, 如果我们对 $M(s)$ 取 n 次 s 的导数, 通过类似的计算有

$$\frac{\mathrm{d}^n}{\mathrm{d}s^n}M(s)\bigg|_{s=0}=\int_{-\infty}^{\infty}x^n f_X(x)\mathrm{d}x=\mathrm{E}[X^n].$$

例： 指数随机变量的概率密度函数为

$$f_X(x)=\lambda\mathrm{e}^{-\lambda x},\quad x\geqslant 0,$$

前面已得

$$M(s)=\frac{\lambda}{\lambda-s}.$$

因此,

$$\frac{\mathrm{d}}{\mathrm{d}s}M(s)=\frac{\lambda}{(\lambda-s)^2},\quad \frac{\mathrm{d}^2}{\mathrm{d}s^2}M(s)=\frac{2\lambda}{(\lambda-s)^3}$$

令 $s=0$, 有

$$\mathrm{E}[X]=\frac{1}{\lambda},\quad \mathrm{E}[X^2]=\frac{2}{\lambda^2},$$

§4 矩母函数

矩母函数的两个更有用且普遍的性质

对于任意的随机变量 X,

$$M_X(0) = \mathrm{E}[\mathrm{e}^{0X}] = \mathrm{E}[1] = 1$$

且如果 X 仅取非负整数值时, 有

$$\lim_{s \to -\infty} M_X(s) = \mathrm{P}(X = 0)$$

§4 矩母函数

4.3 矩母函数的可逆性

矩母函数可逆的条件

假定随机变量 X 的矩母函数 $M_X(s)$ 满足: 存在一个正数 a, 对在区间 $[-a,a]$ 中的任意 s, $M_X(s)$ 都是有限的, 则矩母函数 $M_X(s)$ 唯一地决定 X 的分布函数.

§4 矩母函数

例 4.29 (几何随机变量的矩母函数) 已知随机变量 X 的矩母函数为

$$M(s)=\frac{pe^s}{1-(1-p)e^s},$$

$$\frac{1}{1-\alpha}=1+\alpha+\alpha^2+\cdots,$$

对 $\alpha=(1-p)e^s$ 运用此公式, $(1-p)e^s<1$. 此时, 矩母函数具有展开式

$$M(s)=pe^s(1+(1-p)e^s+(1-p)^2e^{2s}+(1-p)^3e^{3s}+\cdots).$$

概率 $P(X=k)$ 可以通过读取 e^{ks} 的系数得到.

$$P(X=k)=p(1-p)^{k-1},\quad k=1,2,\cdots.$$

§4 矩母函数

4.4 独立随机变量和

独立随机变量的和的矩母函数是和项的矩母函数的乘积.

记 X 和 Y 为独立的随机变量, 并记 $Z = X + Y$.

$$M_Z(s) = \mathrm{E}[\mathrm{e}^{sX}]\mathrm{E}[\mathrm{e}^{sY}] = M_X(s)M_Y(s).$$

如果 $X_1, \cdots, X_n$ 是独立的随机变量, 且 $Z = X_1 + \cdots + X_n$,

$$M_Z(s) = M_{X_1}(s) \cdots M_{X_n}(s).$$

§4 矩母函数

例 4.33 (独立正态随机变量之和仍为正态随机变量) 设 X 和 Y 为两个相互独立的正态随机变量, 均值分别为 μ_x, μ_y, 方差分别为 σ_x^2, σ_y^2. 记 $Z = X + Y$, 则

$$M_X(s) = \mathrm{e}^{\frac{\sigma_x^2 s^2}{2} + \mu_x s}, \qquad M_Y(s) = \mathrm{e}^{\frac{\sigma_y^2 s^2}{2} + \mu_y s},$$

且

$$M_Z(s) = \mathrm{e}^{\frac{(\sigma_x^2 + \sigma_y^2) s^2}{2} + (\mu_x + \mu_y) s}.$$

因此, Z 的矩母函数与均值为 $\mu_x + \mu_y$, 方差为 $\sigma_x^2 + \sigma_y^2$ 的正态随机变量的矩母函数相同.

§4 矩母函数

矩母函数及其性质的小结

- 随机变量 X 的矩母函数定义如下:

$$M_X(s)=\mathrm{E}[\mathrm{e}^{sX}]=\begin{cases}\displaystyle\sum_x \mathrm{e}^{sx}p_X(x), & \text{若 } X \text{ 为离散型},\\ \displaystyle\int_{-\infty}^{\infty}\mathrm{e}^{sx}f_X(x)\mathrm{d}x, & \text{若 } X \text{ 为连续型}.\end{cases}$$

- 随机变量的分布完全由它的矩母函数确定.
- 利用矩母函数计算随机变量的各阶矩:

$$M_X(0)=1,\quad \frac{\mathrm{d}}{\mathrm{d}s}M_X(s)\Big|_{s=0}=\mathrm{E}[X],\quad \frac{\mathrm{d}^n}{\mathrm{d}s^n}M_X(s)\Big|_{s=0}=\mathrm{E}[X^n].$$

- 若 $Y=aX+b$, 则 $M_Y(s)=\mathrm{e}^{sb}M_X(as)$.
- 若 X 和 Y 相互独立, 则 $M_{X+Y}(s)=M_X(s)M_Y(s)$.

§4 矩母函数

常见的离散随机变量的矩母函数

- 参数为 p 的伯努利分布 $(k=0,1)$

$$p_X(k)=\begin{cases}p, & 若\ k=1,\\ 1-p, & 若\ k=0.\end{cases}\qquad M_X(s)=1-p+pe^s.$$

- 参数为 (n,p) 的二项分布， $(k=0,1,\cdots,n)$

$$p_X(k)=\binom{n}{k}p^k(1-p)^{n-k},\qquad M_X(s)=(1-p+pe^s)^n.$$

- 参数为 p 的几何分布 $(k=1,2,\cdots)$

$$p_X(k)=p(1-p)^{k-1},\qquad M_X(s)=\frac{pe^s}{1-(1-p)e^s}.$$

- 泊松分布，参数为 λ， $(k=0,1,\cdots)$

$$p_X(k)=\frac{e^{-\lambda}\lambda^k}{k!},\qquad M_X(s)=e^{\lambda(e^s-1)}.$$

- (a,b) 上的均匀分布 $(k=a,a+1,\cdots,b)$

$$p_X(k)=\frac{1}{b-a+1},\qquad M_X(s)=\frac{e^{as}}{b-a+1}\cdot\frac{e^{(b-a+1)s}-1}{e^s-1}.$$

§ 4 矩母函数

常见连续随机变量的矩母函数

- (a,b) 上的均匀分布 $(a \leqslant x \leqslant b)$

$$f_X(x) = \frac{1}{b-a}, \qquad M_X(s) = \frac{1}{b-a} \cdot \frac{e^{sb} - e^{sa}}{s}.$$

- 参数为 λ 的指数分布 $(x \geqslant 0)$

$$f_X(x) = \lambda e^{-\lambda x}, \qquad M_X(s) = \frac{\lambda}{\lambda - s}, \quad (s < \lambda).$$

- 参数为 (μ, σ^2) 的正态分布 $(-\infty < x < \infty)$

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}, \qquad M_X(s) = e^{(\sigma^2 s^2/2) + \mu s}.$$

§4 矩母函数

4.5 联合分布的矩母函数

考虑同一试验中的 n 个随机变量 $X_1,\cdots,X_n$. 记 $s_1,\cdots,s_n$ 为无量纲实参数. 多元矩母函数是这 n 个参数的函数, 它定义为

$$M_{X_1,\cdots,X_n}(s_1,\cdots,s_n)=\mathrm{E}[\mathrm{e}^{s_1X_1+\cdots+s_nX_n}].$$

如果 $Y_1,\cdots,Y_n$ 是另一组随机变量, 且 $M_{X_1,\cdots,X_n}(s_1,\cdots,s_n)$ 与 $M_{Y_1,\cdots,Y_n}(s_1,\cdots,s_n)$ 相同, 则 $X_1,\cdots,X_n$ 的联合分布与 $Y_1,\cdots,Y_n$ 的联合分布相同.

§5 随机数个相互独立的随机变量之和

5.1 随机变量数目本身随机

随机变量的数目本身也是随机的

$$Y = X_1 + \cdots + X_N$$

- $\mathrm{E}[Y] = \mathrm{E}[X]\mathrm{E}[N]$.
- $\mathrm{var}(Y) = \mathrm{var}(X)\mathrm{E}[N] + (\mathrm{E}[X])^2\mathrm{var}(N)$.
- 矩母函数 $M_Y(s)$ 可由计算矩母函数 $M_N(s)$ 的公式得到，将其中的 e^s 全部替换成 $M_X(s)$ 即可.

§5 随机数个相互独立的随机变量之和

确定某非负整数 n. 随机变量 $X_1+\cdots+X_n$ 与 N 独立. 由此可知, $X_1+\cdots+X_n$ 与事件 $\{N=n\}$ 相互独立. 因此,

$$
\begin{aligned}
\mathrm{E}[Y|N=n] &= \mathrm{E}[X_1+\cdots+X_N|N=n] \\
&= \mathrm{E}[X_1+\cdots+X_n|N=n] \\
&= \mathrm{E}[X_1+\cdots+X_n] \\
&= n\mathrm{E}[X].
\end{aligned}
$$

这对于任意正整数 n 都成立. 因此

$$\mathrm{E}[Y|N] = N\mathrm{E}[X].$$

§5 随机数个相互独立的随机变量之和

使用重期望法则, 有

$$\mathrm{E}[Y] = \mathrm{E}[\mathrm{E}[Y|N]] = \mathrm{E}[N\mathrm{E}[X]] = \mathrm{E}[X]\mathrm{E}[N].$$

类似地,

$$\begin{aligned}\mathrm{var}(Y|N=n) &= \mathrm{var}(X_1+\cdots+X_N|N=n)\\ &= \mathrm{var}(X_1+\cdots+X_n)\\ &= n\mathrm{var}(X).\end{aligned}$$

因为这对任意正整数 n 都是成立的, 随机变量 $\mathrm{var}(Y|N)$ 等于 $N\mathrm{var}(X)$.

§5 随机数个相互独立的随机变量之和

运用全方差法则得

$$\begin{aligned}
\mathrm{var}(Y) &= \mathrm{E}[\mathrm{var}(Y|N)] + \mathrm{var}(\mathrm{E}[Y|N]) \\
&= \mathrm{E}[N\mathrm{var}(X)] + \mathrm{var}(N\mathrm{E}[X]) \\
&= \mathrm{E}[N]\mathrm{var}(X) + (\mathrm{E}[X])^2\mathrm{var}(N).
\end{aligned}$$

基于 $N=n$ 的条件, Y 是独立随机变量 $X_1,\cdots,X_n$ 的和,

$$\begin{aligned}
\mathrm{E}[\mathrm{e}^{sY}|N=n] &= \mathrm{E}[\mathrm{e}^{sX_1}\cdots\mathrm{e}^{sX_N}|N=n] \\
&= \mathrm{E}[\mathrm{e}^{sX_1}\cdots\mathrm{e}^{sX_n}] \\
&= \mathrm{E}[\mathrm{e}^{sX_1}]\cdots\mathrm{E}[\mathrm{e}^{sX_n}] \\
&= (M_X(s))^n,
\end{aligned}$$

§5 随机数个相互独立的随机变量之和

这里 $M_X(s)$ 为 X_i 的矩母函数 (对于任意 i). 运用重期望法则, Y 的 (无条件) 矩母函数为

$$M_Y(s) = \mathrm{E}[\mathrm{e}^{sY}] = \mathrm{E}[\mathrm{E}[\mathrm{e}^{sY}|N]] = \mathrm{E}[(M_X(s))^N] = \sum_{n=1}^{\infty}(M_X(s))^n p_N(n).$$

与下列公式相对照

$$M_N(s) = \mathrm{E}[\mathrm{e}^{sN}] = \sum_{n=1}^{\infty}(\mathrm{e}^s)^n p_N(n),$$

可见 $M_Y(s)$ 和 $M_N(s)$ 形式完全相同, 或者等价地, 将 $M_N(s)$ 的表达式中所有 e^s 用 $M_X(s)$ 替换即可得到 $M_Y(s)$.

§5 随机数个相互独立的随机变量之和

例 4.35 (个数服从几何分布的独立指数随机变量之和) 简为买一本《远大前程》的书逛了很多家书店. 每家书店有这本书的概率都是 p, 且与其他书店相互独立. 逛任意一家书店, 简停留的时间都是随机变量, 服从参数为 λ 的指数分布, 直到她找到这本书或者她肯定这家书店没有这本书后才离开. 假定简会一直逛下去直到她买到这本书, 且她在每家书店停留的时间与其他任何事情都独立. 我们希望求出简逛书店的时间总和的均值、方差和概率密度函数.

§5 随机数个相互独立的随机变量之和

简逛的书店数目 N 服从参数为 p 的几何分布. 因此, 在书店中花费的总时间 Y 是 N 个独立同分布指数随机变量 $X_1, X_2, \cdots, X_N$ 的和, 其中变量 X_i 服从指数分布, 参数为 λ. 我们有

$$\mathrm{E}[Y] = \mathrm{E}[N]\mathrm{E}[X] = \frac{1}{p} \cdot \frac{1}{\lambda}.$$

运用几何分布和指数分布随机变量的方差公式, 得到

$$\operatorname{var}(Y) = \mathrm{E}[N]\operatorname{var}(X) + (\mathrm{E}[X])^2 \operatorname{var}(N) = \frac{1}{p} \cdot \frac{1}{\lambda^2} + \frac{1}{\lambda^2} \cdot \frac{1-p}{p^2} = \frac{1}{\lambda^2 p^2}.$$

为得到矩母函数 $M_Y(s)$, 首先有

$$M_X(s) = \frac{\lambda}{\lambda - s}, \qquad M_N(s) = \frac{pe^s}{1-(1-p)e^s}.$$

§5 随机数个相互独立的随机变量之和

将 $M_N(s)$ 公式中每个 e^s 都换成 $M_X(s)$, 即得

$$M_Y(s) = \frac{pM_X(s)}{1-(1-p)M_X(s)} = \frac{\dfrac{p\lambda}{\lambda - s}}{1-(1-p)\dfrac{\lambda}{\lambda - s}},$$

经过化简可得

$$M_Y(s) = \frac{p\lambda}{p\lambda - s}.$$

这就是服从参数为 $p\lambda$ 的指数随机变量的矩母函数, 所以,

$$f_Y(y) = p\lambda \mathrm{e}^{-p\lambda y}, \qquad y \geqslant 0.$$

作业

29. 设 X 为取值 1, 2, 3 的随机变量, 分布列如下:

$$\mathrm{P}(X=1)=\frac{1}{2},\quad \mathrm{P}(X=2)=\frac{1}{4},\quad \mathrm{P}(X=3)=\frac{1}{4}.$$

求 X 的矩母函数并且用它得到前三个矩, $\mathrm{E}[X], \mathrm{E}[X^2], \mathrm{E}[X^3]$.