概率论笔记和习题

ayhe123

2022 年 6 月 9 日

1 第 1 章笔记

设 T 是 (有限或可列或不可列的) 指标集. 定义

$$\bigcup_{t\in T} A_t := \{x : \exists t \in T, x \in A_t\}, \quad \bigcap_{t\in T} A_t := \{x : \forall t \in T, x \in A_t\}. \tag{1.1}$$

用 "$\exists$" 和 "$\forall$" 来定义无穷集合的并和交是非常重要的观念, 因为对于无穷集合, 很难用直观的方法 (比如 Venn 图) 来理解.

定理 1.1 (De-Morgan)**.** 设 T 是指标集, 则

$$\overline{\bigcup_{t\in T} A_t} = \bigcap_{t\in T} \overline{A_t}, \quad \overline{\bigcap_{t\in T} A_t} = \bigcup_{t\in T} \overline{A_t}.$$

证明. 有

$$\begin{aligned}
\overline{\bigcup_{t\in T} A_t} &= \{x : \neg(\exists t \in T, x \in A_t)\} \\
&= \{x : \forall t \in T, x \notin A_t\} \\
&= \{x : \forall t \in T, x \in \overline{A_t}\} = \bigcap_{t\in T} \overline{A_t},
\end{aligned}$$

$$\begin{aligned}
\overline{\bigcap_{t\in T} A_t} &= \{x : \neg(\forall t \in T, x \in A_t)\} \\
&= \{x : \exists t \in T, x \notin A_t\} \\
&= \{x : \exists t \in T, x \in \overline{A_t}\} = \bigcup_{t\in T} \overline{A_t}.
\end{aligned}$$

□

定理 1.2. 设 T_1, T_2 是指标集, 则

$$\left(\bigcup_{t\in T_1} A_t\right) \cap \left(\bigcup_{u\in T_2} A_u\right) = \bigcup_{\substack{t\in T_1\\u\in T_2}} A_t \cap A_u, \quad \left(\bigcap_{t\in T_1} A_t\right) \cup \left(\bigcap_{u\in T_2} A_u\right) = \bigcap_{\substack{t\in T_1\\u\in T_2}} A_t \cup A_u.$$

证明. 第一个等式的两边都表示集合 $\{\omega : \exists t \in T_1, \exists u \in T_2, \omega \in A_t \wedge \omega \in A_u\}$, 第二个等式的两边都表示集合 $\{\omega : \forall t \in T_1, \forall u \in T_2, \omega \in A_t \vee \omega \in A_u\}$. □

书上的例 4.3 中, 集合序列 $\frac{\#(C \cap B_n)}{n}$ 单调递增且有上界 1, $\therefore \lim\limits_{n\to\infty} \frac{\#(C \cap B_n)}{n}$ 一定存在.

书上的定理 6.1 中, 取交或并的下限不一定是 1, 比如对单调增的序列, $\forall m \in \mathbb{N}_+$ 都有

$$P\left(\bigcup_{j=m}^{\infty} A_j\right) = \lim_{n\to\infty} P(A_n).$$

由概率的连续性可以得到下面的定理:

定理 1.3.

$$P\left(\bigcup_{j=n}^{\infty} A_j\right) = \lim_{m\to\infty} P\left(\bigcup_{j=n}^{m} A_j\right), \quad P\left(\bigcap_{j=n}^{\infty} A_j\right) = \lim_{m\to\infty} P\left(\bigcap_{j=n}^{m} A_j\right).$$

证明. 这里只证明第一个等式, 另一个类似.

设 $C_m = \bigcup\limits_{j=n}^{m} A_j$, 则 C_m 是单调增的. $\therefore$

$$RHS = \lim_{m\to\infty} P(C_m) = P\left(\bigcup_{j=n}^{\infty} C_j\right) = P\left(\bigcup_{j=n}^{\infty} \bigcup_{k=n}^{j} A_k\right) = LHS.$$ □

2 第 1 章习题

1.13. 直径为 1 的硬币随机落在均匀的方格纸上, 问方格的边长为多少才能使硬币和网格不相交的概率小于 0.01.

解. 作硬币的外接正方形 S, 使得 S 的各边与网格平行. 容易看出, 硬币和网格相交当且仅当 S 和网格相交.

$\because S$ 是硬币的外接正方形, $\therefore S$ 的边长为 1.

考虑 S 的任一顶点 (下面以左下角为例) A. S 和网格不相交当且仅当 A 位于图 1 中的蓝色区域.

设方格的边长为 x, 则图 1 中的蓝色区域的边长为 $x - 1$. $\because A$ 是在方格上随机取点, $\therefore A$ 落在蓝色区域的概率

$$p = \frac{(x-1)^2}{x^2} \le 0.01.$$

解得 $\frac{1}{11} \le x \le \frac{10}{9}$.

当 $x \le 1$ 时硬币和网格一定相交, $\therefore$ 条件为 $1 \le x \le \frac{10}{9}$.

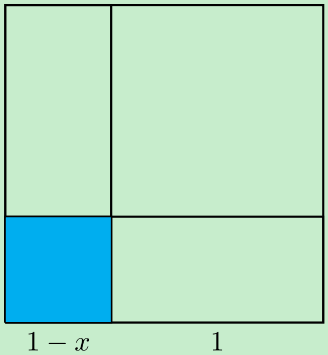

图 1: S 和网格不相交的充要条件

1.14. 在同一小时内有两辆汽车独立地到达同一个加油站加油, 车 A 加油需要 5min, 车 B 加油需要 8min. 如果每辆车在这一小时内等可能地到达, 计算这两辆汽车在加油站不能相遇的概率.

解. 设 A, B 在加油站的时间分别为区间 $(x, x+5)$ 和 $(y, y+8)$, 如图 2.

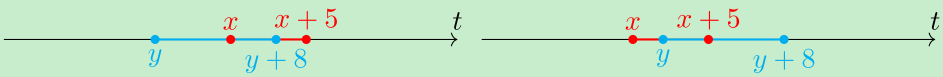

图 2: A, B 在加油站相遇的两种情况, 红色表示 A, 蓝色表示 B

考虑两辆汽车在加油站相遇的情形会比较方便. 从图 2 可以看出要使得两辆汽车在加油站相遇, A, B 到达加油站的时刻 x, y 应满足以下条件[1]:

$$\begin{cases} y+8>x, \\ x+5>y. \end{cases}$$

设 $\Omega=[0,60]^2$, Ω 中元素的第一个分量是 A 到达加油站的时刻, 第二个分量是 B 到达加油站的时刻.

设事件 C 表示两辆汽车在加油站相遇, 则 C 对应的集合为

$$\begin{cases} y+8>x, \\ x+5>y, \\ 0<x<60, \\ 0<y<60. \end{cases}$$

C 对应的区域是图 3 中的蓝色区域, $\overline{C}$ 对应的区域是图 3 左上和右下的两个三角形.

$\therefore$

$$P(\overline{C})=\frac{\dfrac{55^2}{2}+\dfrac{52^2}{2}}{60^2}\approx 0.796.$$

[1] $\because$ 区域的边界的测度为 0, 不会影响概率, $\therefore$ 不用考虑是否取等号.

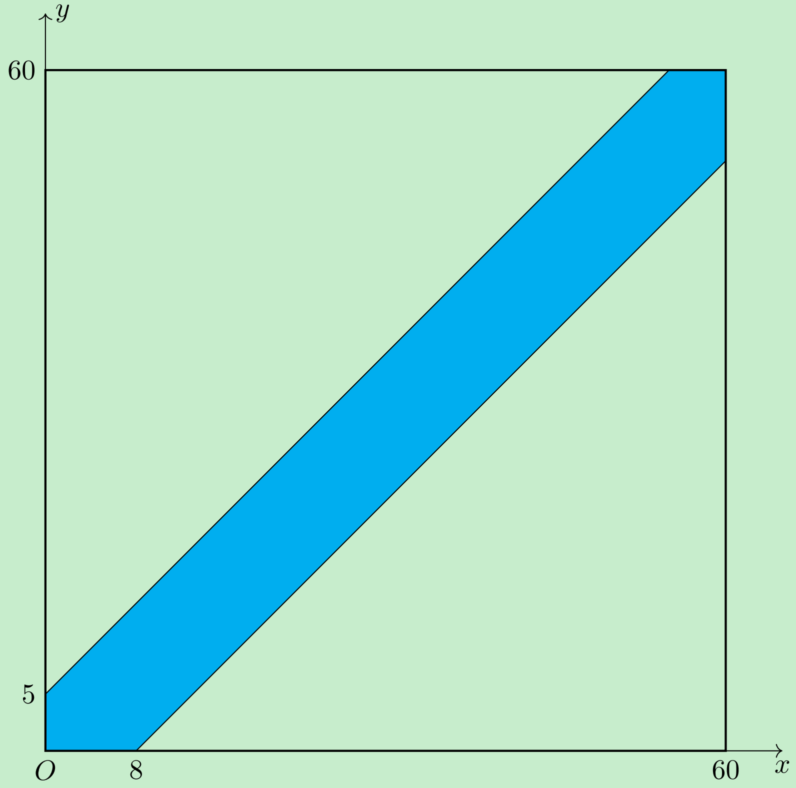

图 3: C 对应的区域

1.17. 证明:

$$P\left(\bigcup_{j=1}^{\infty} A_j\right) \leq \sum_{j=1}^{\infty} P(A_j).$$

证明. 设 $A_0 = \varnothing, B_j = A_j - \bigcup_{i=1}^{j-1} A_i$.

设 $\omega \in \bigcup_{j=1}^{\infty} B_j$. 由式 (1.1) 的定义, $\exists i \in \mathbb{N}_+$ 使得 $\omega \in B_i$.

由 $B_i \subset A_i$ 得 $\omega \in A_i$. $\therefore \omega \in \bigcup_{j=1}^{\infty} B_j$. $\therefore \bigcup_{j=1}^{\infty} B_j \subset \bigcup_{j=1}^{\infty} A_j$.

设 $\omega \in \bigcup_{j=1}^{\infty} A_j$. 由式 (1.1) 的定义, $\exists i \in \mathbb{N}_+$ 使得 $\omega \in A_i$.

$\therefore \exists k \in \mathbb{N}_+$ 使得 $\omega \in A_k$, 且 $\forall i < k, \omega \notin A_i$. $\therefore \omega \notin \bigcup_{i=1}^{k-1} A_i$, $\therefore \omega \in B_k$. $\therefore \bigcup_{j=1}^{\infty} A_j \subset \bigcup_{j=1}^{\infty} B_j$.

$\therefore \bigcup_{j=1}^{\infty} A_j = \bigcup_{j=1}^{\infty} B_j$. $\therefore$

$$P\left(\bigcup_{j=1}^{\infty} A_j\right) = P\left(\bigcup_{j=1}^{\infty} B_j\right).$$

$\because \forall j, k \in \mathbb{N}_+, B_j \subset A_j, B_{j+k} = A_{j+k} - \bigcup_{i=1}^{j+k-1} A_i \subset A_{j+k} - A_j$, $\therefore B_j, B_{j+k}$ 互不相容. 由可列可加性得

$$P\left(\bigcup_{j=1}^{\infty} B_j\right) = \sum_{j=1}^{\infty} P(B_j).$$

$\because B_j \subset A_j$, 由书上的定理 5.1 (5) 得 $P(B_j) < P(A_j)$, $\therefore$

$$\sum_{j=1}^{\infty} P(B_j) \leq \sum_{j=1}^{\infty} P(A_j).$$ □

1.18. 设 $A_1, A_2, \cdots$ 的概率都为 1, 证明:

$$P\left(\bigcap_{j=1}^{n} A_j\right) = 1, \quad P\left(\bigcap_{j=1}^{\infty} A_j\right) = 1.$$

证明. 设 $B_i = \overline{A}_i$, 则 $B_1, B_2, \cdots$ 的概率都为 0.

由书上的例 5.2 得

$$P\left(\bigcup_{j=1}^{n} B_j\right) = 0, \quad P\left(\bigcup_{j=1}^{\infty} B_j\right) = 0.$$

由书上的定理 5.1 (3) 得

$$P\left(\overline{\bigcup_{j=1}^{n} B_j}\right) = 1, \quad P\left(\overline{\bigcup_{j=1}^{\infty} B_j}\right) = 1.$$

由定理 1.1 得

$$P\left(\bigcap_{j=1}^{n} A_j\right) = P\left(\bigcap_{j=1}^{n} \overline{B}_j\right) = P\left(\overline{\bigcup_{j=1}^{n} B_j}\right) = 1,$$

$$P\left(\bigcap_{j=1}^{\infty} A_j\right) = P\left(\bigcap_{j=1}^{\infty} \overline{B}_j\right) = P\left(\overline{\bigcup_{j=1}^{\infty} B_j}\right) = 1. \qquad \square$$

1.19. 对于事件 A, B, 验证

$$P(AB) + P(A\overline{B}) + P(\overline{A}B) + P(\bar{A}\bar{B}) = 1.$$

证明. 由定理 1.2 得

$$\Omega = (A + \overline{A})(B + \overline{B}) = AB + A\overline{B} + \overline{A}B + \bar{A}\bar{B}.$$

容易验证 $AB, A\overline{B}, \overline{A}B, \bar{A}\bar{B}$ 互不相容, 由可列可加性得

$$1 = P(\Omega) = P(AB) + P(A\overline{B}) + P(\overline{A}B) + P(\bar{A}\bar{B}). \qquad \square$$

3 第 2 章笔记

补充几个定理的证明.

定理 3.1. 如果事件 A 的概率为 1 或 0, 则 A 与任意事件 B 相互独立.

证明. 设 $P(A) = 1$, 则由书上第 1 章的例 5.1 得 $P(AB) = P(B) = P(A)P(B)$, 即 A, B 相互独立.

设 $P(A) = 0$, 则 $P(\overline{A}) = 1$, $\therefore \overline{A}, B$ 相互独立, $\therefore A, B$ 相互独立. $\square$

定理 3.2 (Borel-Cantelli 引理)**.** 对事件列 $\{A_n\}$, 有:

(1) 如果 $\sum\limits_{j=1}^{\infty} P(A_j) < \infty$, 则 $P\left(\bigcap\limits_{n=1}^{\infty} \bigcup\limits_{j=n}^{\infty} A_j\right) = 0$;

(2) 如果 $\{A_j\}$ 相互独立, $\sum\limits_{j=1}^{\infty} P(A_j) = \infty$, 则 $P\left(\bigcap\limits_{n=1}^{\infty} \bigcup\limits_{j=n}^{\infty} A_j\right) = 1$.

证明. (1) 令 $C_n = \bigcup\limits_{j=n}^{\infty} A_j$, 则 C_n 是单调减的, 由书上第 1 章的定理 6.1 得

$$P\left(\bigcap_{n=1}^{\infty} \bigcup_{j=n}^{\infty} A_j\right) = P\left(\bigcap_{n=1}^{\infty} C_j\right) = \lim_{n\to\infty} P(C_n).$$

由书上第 1 章的定理 5.1 (6) 得

$$\lim_{n\to\infty} P(C_n) = \lim_{n\to\infty} P\left(\bigcup_{j=n}^{\infty} A_j\right) \le \lim_{n\to\infty} \sum_{j=n}^{\infty} P(A_j).$$

$\because \sum\limits_{j=1}^{\infty} P(A_j) < \infty, \therefore \sum\limits_{j=n}^{\infty} P(A_j) \to 0 (n\to\infty).\ \therefore \lim\limits_{n\to\infty} P(C_n) = 0.$

(2) 与 (1) 一样, 有

$$P\left(\bigcap_{n=1}^{\infty}\bigcup_{j=n}^{\infty} A_j\right) = \lim_{n\to\infty} P\left(\bigcup_{j=n}^{\infty} A_j\right).$$

由定理 1.3 得

$$\lim_{n\to\infty} P\left(\bigcup_{j=n}^{\infty} A_j\right) = \lim_{n\to\infty}\lim_{m\to\infty} P\left(\bigcup_{j=n}^{m} A_j\right).$$

由定理 1.1 得

$$\begin{aligned}\lim_{n\to\infty}\lim_{m\to\infty} P\left(\bigcup_{j=n}^{m} A_j\right) &= \lim_{n\to\infty}\lim_{m\to\infty}\left(1 - P\left(\bigcap_{j=n}^{m}\overline{A}_j\right)\right)\\ &= \lim_{n\to\infty}\left(1 - \lim_{m\to\infty} P\left(\bigcap_{j=n}^{m}\overline{A}_j\right)\right).\end{aligned}$$

由书上的例 2.5 (2) 得 $\overline{A}_j$ 相互独立. $\therefore$

$$\begin{aligned}\lim_{n\to\infty}\left(1 - \lim_{m\to\infty} P\left(\bigcap_{j=n}^{m}\overline{A}_j\right)\right) &= \lim_{n\to\infty}\left(1 - \lim_{m\to\infty}\prod_{j=n}^{m} P(\overline{A}_j)\right)\\ &= \lim_{n\to\infty}\left(1 - \lim_{m\to\infty}\prod_{j=n}^{m}(1 - P(A_j))\right).\end{aligned}$$

$\because \forall x > 0, 1 - x \le e^{-x}, \therefore$

$$\begin{aligned}\lim_{n\to\infty}\left(1 - \lim_{m\to\infty}\prod_{j=n}^{m}(1-P(A_j))\right) &\le \lim_{n\to\infty}\left(1 - \lim_{m\to\infty}\prod_{j=n}^{m}\exp(-P(A_j))\right)\\ &= \lim_{n\to\infty}\left(1 - \lim_{m\to\infty}\exp\left(-\sum_{j=n}^{m} P(A_j)\right)\right)\\ &= \lim_{n\to\infty}\left(1 - \exp\left(-\lim_{m\to\infty}\sum_{j=n}^{m} P(A_j)\right)\right)\\ &= \lim_{n\to\infty}\left(1 - \exp\left(-\sum_{j=n}^{\infty} P(A_j)\right)\right).\end{aligned}$$

$\because \forall n, \sum\limits_{j=n}^{\infty} P(A_j) = \infty, \therefore \forall n, \exp\left(-\sum\limits_{j=n}^{\infty} P(A_j)\right) = 0.\ \therefore$

$$\lim_{n\to\infty}\left(1 - \exp\left(-\sum_{j=n}^{\infty} P(A_j)\right)\right) = \lim_{n\to\infty}(1-0) = 1. \qquad \square$$

在第 (2) 条中有 $\{A_j\}$ 相互独立的条件. 下面的例子说明, 如果 $\{A_j\}$ 不相互独立, 则 Borel-Cantelli 引理不一定成立.

4 第 2 章习题

2.10. 某人参加微积分和线性代数的考试, 无论哪门课先考, 微积分和线性代数通过的概率分别是 a, b, 在微积分先考并通过的条件下, 线性代数通过的概率是 c. 当微积分先考时, 计算

(a) 两门课都通过的概率;

(b) 微积分没通过的条件下, 线性代数通过的概率;

(d) 至少一门通过的概率;

(e) 至少一门不通过的概率.

解. 设 “微积分通过” 和 “线性代数通过” 的事件分别为 A, B, “微积分先考” 的事件为 C, 则 $P_C(A) = P_{\overline{C}}(A) = a, P_C(B) = P_{\overline{C}}(B) = b, P_C(B|A) = P(B|AC) = c$.

(a) “两门课都通过” 的事件为 AB. 有 $P_C(AB) = P_C(B|A)P_C(A) = ac$.

(b) “微积分没通过的条件下, 线性代数通过” 的事件为 $B|\overline{A}$. 由全概率公式,

$$P_C(B) = P_C(B|A)P_C(A) + P_C(B|\overline{A})P_C(\overline{A}) = ac + (1-a)P_C(B|\overline{A}).$$

$\therefore$

$$P_C(B|\overline{A}) = \frac{b-ac}{1-a}.$$

(d) 设 “至少一门通过” 的事件为 C, 则 $\overline{C} = \bar{A}\bar{B}$. 有

$$\begin{aligned} P_C(C) &= 1 - P_C(\bar{A}\bar{B}) \\ &= 1 - P_C(\overline{B}|\overline{A})P_C(\overline{A}) \\ &= 1 - (1 - P_C(B|\overline{A}))P_C(\overline{A}) \\ &= a + b - ac. \end{aligned}$$

2.12. 投 2 个均匀的骰子, 已知至少有一个骰子的点数是 3.

(a) 计算这两个骰子的点数之和为 7 的概率;

(b) 两个骰子的点数之和更有可能是 7 还是 6?

解. (a) 设 “第一个骰子的点数为 i” 的事件为 A_i, “第二个骰子的点数为 i” 的事件为 B_i, 则 “至少有一个骰子的点数是 3” 的事件为 $A_3 \cup B_3$. 由加法公式,

$$P(A_3 \cup B_3) = P(A_3) + P(B_3) - P(A_3B_3) = \frac{1}{6} + \frac{1}{6} - \frac{1}{36} = \frac{11}{36}.$$

“两个骰子的点数之和为 7 的概率” 的事件为 $A_3B_4 + A_4B_3$, 概率为 $\dfrac{2}{36}$.

$\therefore$ 已知至少有一个骰子的点数是 3 的情况下, 两个骰子的点数之和为 7 的概率为 $\dfrac{2/36}{11/36} = \dfrac{2}{11}$.

(b) “两个骰子的点数之和为 6 的概率” 的事件为 A_3B_3, 概率为 $\dfrac{1}{36}$.

$\therefore$ 已知至少有一个骰子的点数是 3 的情况下, 两个骰子的点数之和为 6 的概率为 $\dfrac{1/36}{11/36} = \dfrac{1}{11} \leq \dfrac{2}{11}$. $\therefore$ 两个骰子的点数之和更有可能是 7.

可以用两种不同的方式来看待条件概率, 一种是作为条件的事件 C 改变了样本空间 Ω, 另一种是 C 没有改变 Ω, 但改变了 $\mathcal{F}$ 上的概率 (从原来的 P 变为 P_C). 我们通常采用后一种看法, 但有时前一种看法会更方便.

解 (另一种解法)**.** 用 (a,b) 来表示一次试验, 含义是 "第一个骰子的点数为 a, 第二个骰子的点数为 b." 已知至少有一个骰子的点数是 3 的情况下, 样本空间 $\Omega = \{(1,3),(2,3),\cdots,(6,3),(3,1),(3,2),\cdots,(3,6)\}$, $^{\#}\Omega = 11$(注意 (3,3) 只计算了一次).

"两个骰子的点数之和为 7 的概率" 的事件为 $\{(3,4),(4,3)\}$, $P(\{(3,4),(4,3)\}) = \dfrac{2}{11}$, "两个骰子的点数之和为 6 的概率" 的事件为 $\{(3,3)\}$, $P(\{(3,3)\}) = \dfrac{1}{11}$.

2.19. 甲和另一人下棋, 每局有甲输, 甲赢, 平局三种结果, 每局获胜者得 1 分, 累积多于对手 2 分者获胜. 设每局的结果相互独立, 甲每局获胜的概率为 p, 求甲最终获胜的概率.

解. 设 "甲在某一局获胜", "某一局是平局", "甲在某一局没获胜" 的事件分别为 A_1, A_2, A_3, "某一局前甲比对手多 k 分且甲最终获胜" 的事件为 B_k. 由全概率公式,

$$P(B_k) = P(B_k|A_1)P(A_1) + P(B_k|A_2)P(A_2) + P(B_k|A_3)P(A_3).$$

设 $r = P(A_3)$, $q(k) = P(B_k)$, 则 $q(-2) = 0, q(2) = 1$, 甲最终获胜的概率为 $q(0)$.

$$q(k) = q(k+1)p + q(k)(1-p-r) + q(k-1)r.$$

$\therefore$

$$q(k) = \frac{p}{p+r}q(k+1) + \frac{r}{p+r}q(k-1).$$

代入得

$$\begin{aligned}
q(0) &= \frac{p}{p+r}q(1) + \frac{r}{p+r}q(-1) \\
&= \frac{p}{p+r}\left(\frac{p}{p+r}q(2) + \frac{r}{p+r}q(0)\right) + \frac{r}{p+r}\left(\frac{p}{p+r}q(0) + \frac{r}{p+r}q(-2)\right) \\
&= \frac{p}{p+r}\left(\frac{p}{p+r} + \frac{r}{p+r}q(0)\right) + \frac{r}{p+r}\cdot\frac{p}{p+r}q(0) \\
&= \frac{p^2}{(p+r)^2} + \frac{2pr}{(p+r)^2}q(0).
\end{aligned}$$

解得

$$q(0) = \frac{p^2}{p^2+r^2}.$$

2.20. 用红, 黑两种药治疗同一种疾病, 为了决定用哪种药, 医生将口袋中放入 r 个红球和 b 个黑球后任取一个球, 取到黑球用黑药, 并将球放回, 取到红球用红药, 并将球放回. 如果药物有效就加入一个相同颜色的球. 用 B_i 表示第 i 次取球时取到黑球, 设两种药都一定有效.

(b) 计算 $P(B_2), P(B_3)$;

(c) 计算 $P(B_n)$.

解. (b)

$$\begin{aligned}P(B_2) &= P(B_2|B_1)P(B_1) + P(B_2|\overline{B}_1)P(\overline{B}_1)\\&= \frac{b+1}{r+b+1}\cdot\frac{b}{r+b} + \frac{b}{r+b+1}\cdot\frac{r}{r+b}\\&= \frac{b(r+b+1)}{(r+b+1)(r+b)} = \frac{b}{r+b}.\end{aligned}$$

由第 1.19 题得 $B_1B_2, B_1\overline{B}_2, \overline{B}_1B_2, \overline{B}_1\overline{B}_2$ 是完备事件组. 由全概率公式,

$$\begin{aligned}P(B_3) &= P(B_3|B_1B_2)P(B_1B_2) + P(B_3|B_1\overline{B}_2)P(B_1\overline{B}_2)\\&\quad + P(B_3|\overline{B}_1B_2)P(\overline{B}_1B_2) + P(B_3|\overline{B}_1\overline{B}_2)P(\overline{B}_1\overline{B}_2)\\&= P(B_3|B_1B_2)P(B_2|B_1)P(B_1) + P(B_3|B_1\overline{B}_2)P(\overline{B}_2|B_1)P(B_1)\\&\quad + P(B_3|\overline{B}_1B_2)P(B_2|\overline{B}_1)P(\overline{B}_1) + P(B_3|\overline{B}_1\overline{B}_2)P(\overline{B}_2|\overline{B}_1)P(\overline{B}_1)\\&= \frac{b+2}{r+b+2}\cdot\frac{b+1}{r+b+1}\cdot\frac{b}{r+b} + \frac{b+1}{r+b+1}\cdot\frac{r}{r+b+1}\cdot\frac{b}{r+b}\\&\quad + \frac{b+1}{r+b+2}\cdot\frac{b}{r+b+1}\cdot\frac{r}{r+b} + \frac{b}{r+b+2}\cdot\frac{r+1}{r+b+1}\cdot\frac{r}{r+b}\\&= \frac{(b+2)(b+1)b + 2(b+1)br + (r+1)rb}{(r+b+2)(r+b+1)(r+b)}\\&= b\cdot\frac{b^2+3b+2+2br+2r+r^2+r}{(r+b+2)(r+b+1)(r+b)} = \frac{b}{r+b}.\end{aligned}$$

(c) $P(B_n) = \dfrac{b}{r+b}$. $n=1$ 的情形是显然的.

假设 $P(B_{n-1}) = \dfrac{b}{r+b}$. 如果开始时口袋中有 r' 个红球和 b' 个黑球, 那么第 $n-1$ 次取球时取到黑球的概率为 $\dfrac{b'}{r'+b'}$.

可以认为试验是从第 2 次取球时开始的, 这时第 $n-1$ 次取球时取到黑球对应事件 $B_n|B_1$ 或 $B_n|\overline{B}_1$. 如果第 1 次取球时取到黑球, 那么试验开始时口袋中有 r 个红球和 $b+1$ 个黑球; 如果第 1 次取球时取到红球, 那么试验开始时口袋中有 $r+1$ 个红球和 b 个黑球. ∴

$$P(B_n|B_1) = \frac{b+1}{r+b+1},\quad P(B_n|\overline{B}_1) = \frac{b}{r+b+1}.$$

由全概率公式得

$$\begin{aligned}P(B_n) &= P(B_n|B_1)P(B_1) + P(B_n|\overline{B}_1)P(\overline{B}_1)\\&= \frac{b+1}{r+b+1}\cdot\frac{b}{r+b} + \frac{b}{r+b+1}\cdot\frac{r}{r+b}\\&= \frac{b}{r+b}.\end{aligned}$$

2.28. 口袋中有一个红球和一个黑球, 每次从口袋中有放回地任取一个球, 且第 j 次取球后再加入 j 个红球. 如果一直做下去, 能够取到黑球多少次?

解. 用 A_j 表示第 j 次取球时取到黑球. 第 j 次取球时口袋中有 $2+\sum\limits_{i=1}^{j-1} i = 2+\dfrac{j(j-1)}{2}$ 个球. $\therefore$

$$P(A_j) = \frac{1}{2+\dfrac{j(j-1)}{2}}.$$

$\because \lim\limits_{j\to\infty} \dfrac{P(A_j)}{1/j^2} = 1$ 而 $\sum\limits_{j=1}^{\infty} \dfrac{1}{j^2} < \infty$, $\therefore \sum\limits_{j=1}^{\infty} P(A_j) < \infty$, 由定理 3.2 得 $P\left(\bigcap\limits_{n=1}^{\infty}\bigcup\limits_{j=n}^{\infty} A_j\right) = 0$, 即只有有限次取到黑球.

2.29. 在习题 2.28 中, 若第 j 次取球后再加入 j 个红球和 1 个黑球, 能够取到黑球多少次?

解. 用 A_j 表示第 j 次取球时取到黑球. 第 j 次取球时口袋中有 $2+(j-1)+\sum\limits_{i=1}^{j-1} i = 1+\dfrac{j(j+1)}{2}$ 个球, 其中 $j-1$ 个是黑球. $\therefore$

$$P(A_j) = \frac{j-1}{1+\dfrac{j(j+1)}{2}}.$$

$\because \lim\limits_{j\to\infty} \dfrac{P(A_j)}{1/j} = 1$ 而 $\sum\limits_{j=1}^{\infty} \dfrac{1}{j} = \infty$, $\therefore \sum\limits_{j=1}^{\infty} P(A_j) = \infty$, 由定理 3.2 得 $P\left(\bigcap\limits_{n=1}^{\infty}\bigcup\limits_{j=n}^{\infty} A_j\right) = 1$, 即有无限次取到黑球.

2.30. 证明全概率公式: 如果事件 $A_1, A_2, \cdots$ 互不相容, $B \subset \bigcup\limits_{j=1}^{\infty} A_j$, 则

$$P(B) = \sum_{j=1}^{\infty} P(A_j)P(B|A_j).$$

证明. 设 $\omega \in B \cap \left(\bigcup\limits_{j=1}^{\infty} A_j\right)$, 则 $\omega \in B$ 且 $\exists j \in \mathbb{N}, \omega \in A_j$. $\therefore \exists j \in \mathbb{N}, \omega \in A_j B$. $\therefore B \cap \left(\bigcup\limits_{j=1}^{\infty} A_j\right) = \bigcup\limits_{j=1}^{\infty} A_j B$

由定理 1.2 得

$$B \cap \left(\bigcup_{j=1}^{\infty} A_j\right) = \bigcup_{j=1}^{\infty} A_j B.$$

$\because A_1, A_2, \cdots$ 互不相容, $\therefore A_1B, A_2B, \cdots$ 互不相容. 由书上第 1 章的定义 4.1(c) 得

$$P(B) = P\left(\bigcup_{j=1}^{\infty} A_j B\right) = \sum_{j=1}^{\infty} P(A_j B).$$

由乘法公式得

$$\sum_{j=1}^{\infty} P(A_j B) = \sum_{j=1}^{\infty} P(B|A_j)P(A_j).$$

□

2.31. 证明 Bayes 公式: 如果事件 $A_1, A_2, \cdots$ 互不相容, $B \subset \bigcup_{j=1}^{\infty} A_j$, 则 $P(B) > 0$ 时有

$$P(A_j|B) = \frac{P(A_j)P(B|A_j)}{\sum_{i=1}^{\infty} P(A_i)P(B|A_i)}, \quad j \geq 1.$$

证明. 由乘法公式得

$$P(A_j|B) = \frac{P(A_j B)}{P(B)} = \frac{P(B|A_j)P(A_j)}{P(B)}.$$

由第 2.30 得

$$\frac{P(B|A_j)P(A_j)}{P(B)} = \frac{P(B|A_j)P(A_j)}{\sum_{i=1}^{\infty} P(A_i)P(B|A_i)}.$$

□

2.33. 如果你的水平略高于对手, 为了保证比赛获胜, 你期望比赛规则是三局两胜还是五局三胜?

解. 见 http://lanqi.org/everyday/19943/.

2.35. 设想如下的抽奖节目: 三扇关闭的门后各有一个奖品, 其中之一是汽车, 其余是羊. 每扇门后是汽车的概率相同. 猜奖者任选一扇门后得到门后的奖品. 当猜奖者选中一扇门尚未打开时, 主持人打开了另外两扇门之一, 发现门后是羊. 这时猜奖者有机会改猜剩下的那扇门. 在下面的情况下, 计算猜奖者换门或不换门得到汽车的概率:

(a) 假设主持人知道汽车在哪扇门后;

(b) 假设主持人不知道汽车在哪扇门后;

(c) 假设主持人知道汽车在哪扇门后的概率为 0.6.

解. 称猜奖者一开始选中的门是 “1 号门”, 主持人打开的门是 “2 号门”, 剩下的那扇门是 “3 号门”, A_i $(i = 1, 2, 3)$ 表示事件 “汽车在 i 号门后”, B 表示事件 “主持人打开 2 号门, 发现门后是羊”. 要求 $P(A_1|B)$ 和 $P(A_3|B)$. 由 Bayes 公式得

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_{j=1}^{3} P(B|A_j)P(A_j)}.$$

$\because P(A_1) = P(A_2) = P(A_3) = \frac{1}{3}$, $\therefore$

$$P(A_i|B) = \frac{P(B|A_i)}{\sum_{j=1}^{3} P(B|A_j)}.$$

首先, “主持人打开 2 号门, 发现门后是羊” 意味着车不可能在 2 号门后, 即 $P(B|A_2)=0$. 其余的条件概率会随情况而变化.

(a) 如果汽车在 1 号门后, 那么主持人可以从 2 号和 3 号门中任选一扇打开, $\therefore P(B|A_1)=\frac{1}{2}$. 如果汽车在 3 号门后, 那么主持人一定会打开 2 号门, $\therefore P(B|A_3)=1$. $\therefore$

$$P(A_1|B)=\frac{\frac{1}{2}}{\frac{1}{2}+1}=\frac{1}{3},\quad P(A_3|B)=\frac{1}{\frac{1}{2}+1}=\frac{2}{3}.$$

(b) $\because$ 主持人不知道汽车在哪扇门后, $\therefore$ 无论汽车在哪扇门后, 主持人都会从 2 号和 3 号门中任选一扇打开, $\therefore P(B|A_1)=P(B|A_3)=\frac{1}{2}$. $\therefore P(A_1|B)=P(A_3|B)=\frac{1}{2}$.

(c) 设 “主持人知道汽车在哪扇门后” 为事件 C, 则由 (a) 得 $P_C(B|A_1)=\frac{1}{2}, P_C(B|A_3)=1$, 由 (b) 得 $P_{\overline{C}}(B|A_1)=P_{\overline{C}}(B|A_3)=\frac{1}{2}$. 由全概率公式得

$$P(B|A_1)=P_C(B|A_1)P(C)+P_{\overline{C}}(B|A_1)P(\overline{C})=\frac{1}{2},$$

$$P(B|A_3)=P_C(B|A_3)P(C)+P_{\overline{C}}(B|A_3)P(\overline{C})=\frac{4}{5}.$$

$\therefore$

$$P(A_1|B)=\frac{\frac{1}{2}}{\frac{1}{2}+\frac{4}{5}}=\frac{5}{13},\quad P(A_3|B)=\frac{\frac{4}{5}}{\frac{1}{2}+\frac{4}{5}}=\frac{8}{13}.$$

2.36. 在书上的例 1.4 中, 当全班有 n 个人时, 用 B_n 表示至少有一人得到自己的作业本, 用 D_k 表示恰好有 k 个人得到自己的作业本, 计算 $^{\#}\overline{B}_n$ 和 $P(D_k)$.

解. 由书上的例 1.4 得 $P(B_n)=\sum\limits_{k=1}^{n}(-1)^{k-1}\frac{1}{k!}$. $\therefore$

$$P(\overline{B}_n)=1-P(B_n)=\sum_{i=0}^{n}(-1)^i\frac{1}{i!}.$$

$\therefore$

$$^{\#}\overline{B}_n={}^{\#}\Omega P(\overline{B}_n)=n!\sum_{i=0}^{n}(-1)^i\frac{1}{i!}.$$

$\overline{B}_{n-k}$ 表示 “把 $n-k$ 个作业本随机分给 $n-k$ 个人, 没有人得到自己的作业本”. 恰好有 k 个人得到自己的作业本意味着剩下 $n-k$ 个人中没有人得到自己的作业本. $\therefore$

$$D_k=\sum_{1\le j_1<j_2<\cdots<j_k\le n}A_{j_1}A_{j_2}\cdots A_{j_k}\overline{B}_{n-k},$$

$$P(D_k)=C_n^k\frac{(n-k)!}{n!}P(\overline{B}_{n-k})=\frac{1}{k!}\sum_{i=0}^{n-k}(-1)^i\frac{1}{i!}.$$

5 第 3 章笔记

为了引入随机变量的定义, 首先需要介绍一些测度论的概念.

定义 5.1. 设 Ω 是任意的集合, $\mathcal{F}$ 是 Ω 上的 σ 域 ($\mathcal{F}$ 是 Ω 的子集构成的集合且 $\mathcal{F}$ 满足书上第 1 章第 4 节开头的条件), 则称 $(\Omega, \mathcal{F})$ 是一个**可测空间**.

定义 5.2. 设 $(\Omega_1, \mathcal{F}_1), (\Omega_2, \mathcal{F}_2)$ 是两个可测空间, 如果函数 $f : \Omega_1 \to \Omega_2$ 满足: $\forall A \in \mathcal{F}_2$, f 在 A 上的原像 $\{\omega | f(\omega) \in A\} \in \mathcal{F}_1$, 则称 f 是一个**可测变换**.

从概率空间 $(\Omega, \mathcal{F}, P)$ 可以自然地得到一个可测空间 $(\Omega, \mathcal{F})$. 设 $\mathcal{R}$ 是包含 $\mathbb{R}$ 上的区间全体的 σ 域. 按照定义, $(\mathbb{R}, \mathcal{R})$ 是一个可测空间, 称为 **Borel 可测集**. 有:

定义 5.3. 设 $(\Omega, \mathcal{F}, P)$ 是概率空间, $(\mathbb{R}, \mathcal{R})$ 是 Borel 可测集, 称可测变换 $X : \Omega \to \mathbb{R}$ 是 Ω 上的**随机变量**.

在初等概率论中, 可以简单地认为 Ω 上的随机变量是 Ω 上的实值函数.

例 5.1 (书上的例 1.3(2))**.** 设 X 是随机变量, $x \in \mathbb{R}$, 则

$$\bigcup_{n=1}^{\infty} \{X \leq x - 1/n\} = \{X < x\}.$$

证明. 设 $\omega \in \bigcup\limits_{n=1}^{\infty} \{X \leq x - 1/n\}$, 则 $\exists m \in \mathbb{N}_+, X(\omega) \leq x - \dfrac{1}{m} < x$. $\therefore \omega \in \{X < x\}$. $\therefore \bigcup\limits_{n=1}^{\infty} \{X \leq x - 1/n\} \subset \{X < x\}$.

设 $\omega \in \{X < x\}$, 则 $\exists \varepsilon > 0$ 使得 $X(\omega) \leq x - \varepsilon$. 取 $m \geq \dfrac{1}{\varepsilon}$, 则 $\omega \in \{X \leq x - 1/m\}$. $\therefore \omega \in \bigcup\limits_{n=1}^{\infty} \{X \leq x - 1/n\}$. $\therefore \bigcup\limits_{n=1}^{\infty} \{X \leq x - 1/n\} \supset \{X < x\}$. □

随机变量 X 可以由其分布函数 $F(x) = P(X \leq x)$ 唯一确定. 之后用 $X \sim F(x)$ 来表示由 $F(x)$ 确定的随机变量.

6 第 3 章习题

3.1. 证明:

$$\lim_{x \to \infty} F(x) = 1, \quad \lim_{x \to -\infty} F(x) = 0.$$

证明. $\because \lim\limits_{x \to \infty} F(x)$ 存在, $\therefore \lim\limits_{x \to \infty} F(x) = \lim\limits_{n \to \infty} F(n) = \lim\limits_{n \to \infty} P(X \leq n)$.

$\because \{X \leq n\}$ 单调递增, 由概率的连续性得

$$\lim_{n \to \infty} P(X \leq n) = P\left(\bigcup_{n=1}^{\infty} \{X \leq n\}\right) = P(X < \infty) = 1.$$

$\because \lim\limits_{x\to-\infty} F(x)$ 存在, $\therefore \lim\limits_{x\to-\infty} F(x) = \lim\limits_{n\to\infty} F(-n) = \lim\limits_{n\to\infty} P(X \le -n)$.
$\because \{X \le -n\}$ 单调递减, 由概率的连续性得

$$\lim_{n\to\infty} P(X \le -n) = P\left(\bigcap_{n=1}^{\infty}\{X \le -n\}\right) = P(X \le -\infty) = 0. \qquad \square$$

3.2. 证明: 如果常数 a, b, c 使得 $g(x) = \exp(ax^2 + bx + c), x \in \mathbb{R}$ 是概率密度, 则 $a < 0$.

证明. 证明逆否命题. 若 $a = 0$, 则 $g(x) = e^{bx+c}$, $\because g$ 在 $\mathbb{R}$ 上不可积, $\therefore g$ 不是概率密度.

若 $a > 0$, 则 $g(x) = \exp\left(a\left(x + \dfrac{b}{2a}\right)^2 - \dfrac{b^2}{4a^2} + c\right)$. $\because g(x) \to \infty$ $(x \to \pm\infty)$, $\therefore g$ 在 $\mathbb{R}$ 上不可积, $\therefore g$ 不是概率密度. $\square$

3.9. 甲每天收到的电子邮件数服从 Poisson 分布 $\mathcal{P}(\lambda)$, 且每封电子邮件被过滤掉的概率为 0.2.

(c) 已知甲看到了自己的 k 封电子邮件, 计算他有 m 封电子邮件被过滤掉的概率.

(d) 甲每天看到的电子邮件数与被过滤掉的电子邮件数是否独立?

解. (c) 设 A_k 表示事件 "甲看到了 k 封电子邮件", B_m 表示事件 "甲有 m 封电子邮件被过滤掉", 则事件 A_kB_m 意味着甲一共收到了 $k+m$ 封电子邮件. $\therefore$

$$P(A_kB_m) = \frac{\lambda^{k+m}}{(k+m)!}e^{-\lambda} \cdot C_{k+m}^k \cdot 0.8^k \cdot 0.2^m.$$

由书上的例 2.3 得甲每天看到的电子邮件数服从 Poisson 分布 $\mathcal{P}(0.8\lambda)$. 有

$$\begin{aligned} P(B_m|A_k) &= \frac{P(A_kB_m)}{P(A_k)} \\ &= \frac{\dfrac{\lambda^{k+m}}{(k+m)!}e^{-\lambda} \cdot C_{k+m}^k \cdot 0.8^k \cdot 0.2^m}{\dfrac{(0.8\lambda)^k}{k!}e^{-0.8\lambda}} \\ &= \frac{\lambda^m}{m!}e^{-0.2\lambda} \cdot 0.2^m. \end{aligned}$$

(d) 由书上的例 2.3 得甲每天被过滤掉的电子邮件数服从 Poisson 分布 $\mathcal{P}(0.2\lambda)$. 有

$$P(B_m) = \frac{(0.2\lambda)^m}{m!}e^{-0.2\lambda} = P(B_m|A_k).$$

$\therefore$ 甲每天看到的电子邮件数与被过滤掉的电子邮件数相互独立.

3.10. 设车间有 100 台型号相同的机床相互独立地工作, 每台机床在时间 $(0, t]$ 内发生故障的概率为 0.01, 发生故障的机床需要一人来维修, 且一人在 $(0, t]$ 内只能维修一台机床. 考虑两种配备维修工人的方法:

(a) 5 个人每人负责 20 台机床.

(b) 3 个人同时负责 100 台机床.

在以上两种情况下计算机床在时间 $(0,t]$ 内发生故障时不能被及时维修的概率.

解. (a) 考虑 5 个维修工人中的任意一个 (记为工人甲). 工人甲负责的机床中发生故障的机床数 n 服从二项分布 $\mathcal{B}(20,0.01)$, 他能及时维修这 20 台机床的概率为 $P(n\le 1)=0.99^{20}+C_{20}^1\cdot 0.99^{19}\cdot 0.01\approx 0.98314$.

$\therefore$ 全部 100 台机床在时间 $(0,t]$ 内发生故障时不能被及时维修的概率为 $1-0.98314^5\approx 0.0815$.

(b) 100 台机床中发生故障的机床数 n 服从二项分布 $\mathcal{B}(100,0.01)$. 机床在时间 $(0,t]$ 内发生故障时不能被及时维修的概率为

$$\begin{aligned}1-P(n\le 3)&=1-(0.99^{100}+C_{100}^1\cdot 0.99^{99}\cdot 0.01+C_{100}^2\cdot 0.99^{98}\cdot 0.01^2+C_{100}^3\cdot 0.99^{97}\cdot 0.01^3)\\&\approx 0.0184.\end{aligned}$$

3.14. 一个使用了 t 小时的电阻在 Δt 内失效的概率是 $\lambda\Delta t+o(\Delta t)$, 设该电阻的使用寿命是连续型随机变量 X, 求 X 的分布.

解. 该电阻在 Δt 内不失效的概率是 $1-\lambda\Delta t-o(\Delta t)$, $\therefore$

$$P(X>t+\Delta t|X>t)=1-\lambda\Delta t-o(\Delta t).$$

$\therefore$

$$P(X>t+\Delta t)=(1-\lambda\Delta t-o(\Delta t))P(X>t).$$

设 X 的分布函数为 $F(x)$, 则

$$\begin{aligned}F(t+\Delta t)&=1-P(X>t+\Delta t)\\&=1-(1-\lambda\Delta t-o(\Delta t))P(X>t)\\&=1-(1-\lambda\Delta t-o(\Delta t))(1-F(t))\\&=(\lambda\Delta t+o(\Delta t))(1-F(t))+F(t).\end{aligned}$$

$\therefore$

$$\frac{F(t+\Delta t)-F(t)}{\Delta t}=(\lambda+o(1))(1-F(t))\quad(\Delta t\to 0).$$

$\therefore$

$$\frac{\mathrm{d}F}{\mathrm{d}t}=\lambda(1-F).$$

解得

$$F(x)=1-Ce^{-\lambda x}.$$

由 $\int_0^\infty F(x)\mathrm{d}x=1$ 得 $C=1$. $\therefore X\sim\mathcal{E}(\lambda)$.

3.16. 设 $X\sim\mathcal{P}(\lambda)$, 计算 $Y=\sqrt{X}$ 的概率分布.

解. $\because X$ 是取值为 $\mathbb{N}$ 的离散型随机变量, $\therefore Y$ 是离散型随机变量. 有

$$P(Y=y)=P(X=y^2)=\frac{\lambda^{y^2}}{(y^2)!}e^{-\lambda},\quad y\in\{z:z^2\in\mathbb{N}\}.$$

3.17. 设电流 I 在 $8\sim 9\mathrm{A}$ 间均匀分布. 当电流通过 2Ω 的电阻时, 消耗的功率 (单位: W) 是 $W=2I^2$. 求 W 的概率密度.

解. 有 $P(128<W\le 162)=P(8<I\le 9)=1$. 对 $w\in(128,162)$,

$$
\begin{aligned}
P(W=w) &= P(2I^2=w)\\
&= P\left(I=\sqrt{\frac{w}{2}}\right)+P\left(I=-\sqrt{\frac{w}{2}}\right)\\
&= 1\cdot\left|\frac{\mathrm{d}}{\mathrm{d}w}\sqrt{\frac{w}{2}}\right|\mathrm{d}w+0\\
&= \frac{1}{\sqrt{8w}}.
\end{aligned}
$$

$\therefore$

$$P_W(w)=\frac{1}{\sqrt{8w}},\quad w\in(128,162).$$

3.20 (c)**.** 设 X 有分段连续的概率密度 $f(x)$. 求 $Y=\tan X$ 的概率密度.

解. Y 的取值为 $\mathbb{R}$. 有

$$
\begin{aligned}
P(Y=y) &= P(\tan X=y)\\
&= \sum_{n=\infty}^{\infty}P(X=\arctan y+n\pi)\\
&= \sum_{n=\infty}^{\infty}\frac{1}{1+y^2}f(\arctan y+n\pi).
\end{aligned}
$$

$\therefore$

$$P_Y(y)=\frac{1}{1+y^2}\sum_{n=\infty}^{\infty}f(\arctan y+n\pi).$$

3.21. 设 $X\sim\mathcal{B}(n,p)$.

(a) 已知 $n=19,p=0.7$, 求 $p_k=P(X=k)$ 的最大值点 k;

(b) 已知 $n=19,X=9$, 求使得 $P(X=9)$ 最大的 p.

解. (a) 设 $k<19$. $\because$

$$P(X=k)=C_{19}^k\cdot 0.7^k\cdot 0.3^{19-k}=\frac{19!}{k!(19-k!)}\cdot 0.7^k\cdot 0.3^{19-k},$$

$$P(X=k+1)=\frac{19!}{(k+1)!(19-k-1!)}\cdot 0.7^{k+1}\cdot 0.3^{19-k-1},$$

$\therefore$

$$\begin{aligned} p_k - p_{k+1} &= \frac{19!}{k!(19-k!)} \cdot 0.7^k \cdot 0.3^{19-k} - \frac{19!}{(k+1)!(19-k-1!)} \cdot 0.7^{k+1} \cdot 0.3^{19-k-1} \\ &= \frac{19!}{k!(19-k-1!)} \cdot 0.7^k \cdot 0.3^{19-k-1} \left(\frac{0.3}{19-k} - \frac{0.7}{k+1} \right). \end{aligned}$$

$\therefore$ 当 $\dfrac{0.3}{19-k} - \dfrac{0.7}{k+1} > 0, k > 13$ 时 $p_k > p_{k+1}$, 当 $k < 13$ 时 $p_k < p_{k+1}$, 当 $k = 13$ 时 $p_k = p_{k+1}$.
$\therefore$ 当 $k = 13, 14$ 时 p_k 取最大值.

(b) $P(X=9) = C_{19}^9 p^9 (1-p)^{10}$. 只需求 $f(p) = \ln P(X=9)$ 的最大值点. 有

$$f(p) = \ln C_{19}^9 + 9 \ln p + 10 \ln(1-p), \quad \frac{\mathrm{d}f}{\mathrm{d}p} = \frac{9}{p} - \frac{10}{1-p}.$$

$\therefore$ 当 $p = \dfrac{9}{19}$ 时 $P(X=9)$ 最大.

3.23. 设 $X \sim \mathcal{P}(\lambda)$.

(a) 已知 $\lambda = 23.8$, 求 $p_k = P(X=k)$ 的最大值点 k;

(b) 已知 $X = 21$, 求使得 $P(X=21)$ 最大的 λ.

解. (a) 对 $k \geq 0$, 有

$$\begin{aligned} P(X=k+1) - P(X=k) &= \frac{\lambda^{k+1}}{(k+1)!} e^{-\lambda} - \frac{\lambda^k}{k!} e^{-\lambda} \\ &= \frac{\lambda^k}{k!} e^{-\lambda} \left(\frac{\lambda}{k+1} - 1 \right). \end{aligned}$$

当 $k < \lambda - 1$ 时有 $P(X=k+1) > P(X=k)$, 当 $k > \lambda - 1$ 时有 $P(X=k+1) < P(X=k)$, $\therefore$ 当 $k = 23$ 时 p_k 取最大值.

(b) $P(X=21) = \dfrac{\lambda^{21}}{21!} e^{-\lambda}$. 令

$$f(\lambda) = \ln P(X=21) = 21 \ln \lambda - \ln 21! - \lambda,$$

有

$$\frac{\mathrm{d}f}{\mathrm{d}\lambda} = \frac{21}{\lambda} - 1.$$

$\therefore$ 当 $\lambda = 21$ 时 $P(X=9)$ 最大.

3.25. 将一个骰子投掷 n 次, 用 m 表示掷得的最小点数, 用 M 表示掷得的最大点数. 计算

(a) $P(m=k), \ 1 \leq k \leq 6$;

(b) $P(M=k), \ 1 \leq k \leq 6$;

(c) $P(m=2, M=5)$.

解. 设第 i 次掷得的点数为 X_i, 则 X_i 相互独立, $P(X_i=k)=\dfrac{1}{6}\ (k=1,\cdots,6)$.

(a)

$$\begin{aligned}P(m\geq k)&=P(X_1\geq k,X_2\geq k,\cdots,X_n\geq k)\\&=P(X_1\geq k)P(X_2\geq k)\cdots P(X_n\geq k)\\&=\left(\frac{6-k+1}{6}\right)^n,\end{aligned}$$

$$P(m>k)=P(m\geq k+1)=\left(\frac{6-(k+1)+1}{6}\right)^n=\left(\frac{6-k}{6}\right)^n,$$

$\therefore$

$$P(m=k)=P(m\geq k)-P(m>k)=\frac{(6-k+1)^n-(6-k)^n}{6^n}.$$

(b)

$$\begin{aligned}P(M\leq k)&=P(X_1\leq k,X_2\leq k,\cdots,X_n\leq k)\\&=P(X_1\leq k)P(X_2\leq k)\cdots P(X_n\leq k)\\&=\left(\frac{k}{6}\right)^n,\end{aligned}$$

$$P(M=k)=P(M\leq k)-P(M\leq k-1)=\frac{k^n-(k-1)^n}{6^n}.$$

(c)

$$P(m\geq k_1,M\leq k_2)=\left(\frac{k_2-k_1+1}{6}\right)^n,\quad k_1\leq k_2.$$

$$\begin{aligned}P(m=2,M=5)&=P(m\geq 2,M\leq 5)-P(\{m\geq 3,M\leq 5\}\vee\{m\geq 2,M\leq 4\})\\&=P(m\geq 2,M\leq 5)-P(m\geq 3,M\leq 5)-P(m\geq 2,M\leq 4)+P(m\geq 3,M\leq 4)\\&=\left(\frac{5-2+1}{6}\right)^n-\left(\frac{5-3+1}{6}\right)^n-\left(\frac{4-2+1}{6}\right)^n+\left(\frac{4-3+1}{6}\right)^n\\&=\frac{2^n}{3^n}-2\frac{1}{2^n}+\frac{1}{3^n}.\end{aligned}$$

3.26. 设点随机地落在中心为原点, 半径为 R 的圆周上, 求落点横坐标的概率密度 $f(x)$.

解. X 的取值为 $(-R,R)$. 当 $0\leq x<R$ 时有

$$\begin{aligned}P(X\leq x)&=\frac{\frac{1}{2}\left(2\pi-2\arccos\frac{x}{R}\right)R^2-2\cdot\frac{x\sqrt{R^2-x^2}}{2}}{\pi R^2}\\&=1-\frac{\arccos\frac{x}{R}}{\pi}-\frac{1}{\pi}\cdot\frac{x}{R}\sqrt{1-\left(\frac{x}{R}\right)^2}.\end{aligned}$$

$\therefore$

$$\begin{aligned} f(x) &= \frac{\partial P(X \le x)}{\partial x} \\ &= \frac{1}{R} \cdot \frac{\partial P(X \le x)}{\partial (x/R)} \\ &= \frac{1}{R} \cdot \left(\frac{2\sqrt{1-(x/R)^2}}{\pi} \right) \\ &= \frac{2\sqrt{R^2-x^2}}{\pi R^2}. \end{aligned}$$

显然 $f(x) = f(-x)$. $\therefore$

$$f(x) = \frac{2\sqrt{R^2-x^2}}{\pi R^2}, \quad -R < x < R.$$

注. 书上的答案应该有点问题. 下面这行 Mathematica 代码可以用来画 $\frac{2\sqrt{16-x^2}}{16\pi}$ (这里的答案当 $R=4$ 时的情形) 以及 $\frac{1}{\pi\sqrt{16-x^2}}$ (书上的答案当 $R=4$ 时的情形) 的图象, 可以看到 $\frac{1}{\pi\sqrt{16-x^2}}$ 的值在 $x = \pm 4$ 处比较大, 在 $x=0$ 处比较小, 这与实际情况不符.

```
Plot[{2 Sqrt[16 - x^2]/(16 Pi), 1/(Pi Sqrt[16 - x^2])}, {x, -4, 4},
  PlotLabels -> Automatic]
```

3.28. 设 X 有概率密度 $f(x) = cx/\pi^2, x \in (0, \pi)$. 求 $Y = \sin X$ 的概率密度.

解. 由 $\int_0^\pi f(x) = 1$ 得 $c = 2$.

Y 的取值为 $(0,1]$. 对 $y \in (0,1]$, 有

$$\begin{aligned} P(Y=y) &= P(\sin X = y) \\ &= P(X = \arcsin y) + P(X = \pi - \arcsin y) \\ &= \frac{2\arcsin y}{\pi^2} \cdot \frac{1}{\sqrt{1-y^2}} + \frac{2(\pi - \arcsin y)}{\pi^2} \cdot \left| -\frac{1}{\sqrt{1-y^2}} \right| \\ &= \frac{2}{\pi} \cdot \frac{1}{\sqrt{1-y^2}}. \end{aligned}$$

3.30. 设 $f(x)$ 是 $[0, \infty)$ 上的连续函数. 证明:

(1) 如果 $\forall x, y > 0$, 有 $f(x+y) = f(x) + f(y)$, 则 $\exists a$ 使得 $f(x) = ax, x \ge 0$;

(2) 如果 $\forall x, y > 0$, 有 $f(x+y) = f(x)f(y) > 0$, 则 $\exists b$ 使得 $f(x) = e^{ax}, x \ge 0$.

证明. (1) 用数学归纳法证明: $\forall n \in \mathbb{N}, f(n) = f(1) \cdot n$. 当 $n = 0$ 时有 $f(0) = f(0+0) = f(0) + f(0)$, $\therefore f(0) = 0 = f(1) \cdot 0$.

假设有 $f(n-1) = f(1) \cdot (n-1)$, 则

$$f(n) = f((n-1)+1) = f(n-1) + f(1) = f(1) \cdot (n-1) + f(1) = nf(1).$$

$\therefore \forall n \in \mathbb{N}, f(n) = f(1) \cdot n.$

设 $a = \dfrac{p}{q} \in \mathbb{Q}_+, p \in \mathbb{N}, Q \in \mathbb{N}_+$. 有

$$\begin{aligned} qf(a) &= f(a) + f(a) + \underbrace{f(a) + \cdots + f(a)}_{q-2\text{个}f(a)} \\ &= f(2a) + \underbrace{f(a) + \cdots + f(a)}_{q-2\text{个}f(a)} \\ &= \cdots \\ &= f(qa) = f(p). \end{aligned}$$

$\because p \in \mathbb{N}, \therefore f(p) = pf(1)$. $\therefore qf(a) = pf(1), f(a) = \dfrac{p}{q} f(1)$.

$\because f$ 是 $[0, \infty)$ 上的连续函数, $\therefore f$ 在 $[0, \infty)$ 上任意一点处的极限存在且等于 f 在该点的值. $\therefore$ 对收敛到 $[0, \infty)$ 上任意一点 x_0 的数列 $a_1, a_2, \cdots$, 有

$$f(x_0) = \lim_{x \to x_0} f(x) = \lim_{n \to \infty} f(a_n).$$

$\forall x \in [0, \infty)$, 存在有理数列 $a_1, a_2, \cdots$ 使得 $\lim_{n \to \infty} a_n = x$, 有

$$\lim_{n \to \infty} f(a_n) = \lim_{n \to \infty} a_n f(1) = f(1) \lim_{n \to \infty} a_n = xf(1).$$

(2) 用数学归纳法证明: $\forall n \in \mathbb{N}, f(n) = (f(1))^n$. 当 $n = 0$ 时有 $f(0) = f(0+0) = f(0) \cdot f(0)$, $\therefore f(0) = 1 = (f(1))^0$.

假设有 $f(n-1) = (f(1))^{n-1}$, 则

$$f(n) = f((n-1)+1) = f(n-1)f(1) = (f(1))^{n-1} f(1) = (f(1))^n.$$

$\therefore \forall n \in \mathbb{N}, f(n) = (f(1))^n.$

设 $a = \dfrac{p}{q} \in \mathbb{Q}_+, p \in \mathbb{N}, Q \in \mathbb{N}_+$. 与 (1) 类似, 有

$$(f(a))^q = f(qa) = f(p) = (f(1))^p.$$

$\therefore$

$$f(a) = (f(1))^{p/q}.$$

与 (1) 类似, 用 f 的连续性可以得到: $\forall x \in [0, \infty)$, $f(x) = (f(1))^x = e^{x \ln f(1)}$. □

7 第 4 章笔记

7.1 边缘分布函数

书上没有给出 n 维随机向量的边缘分布函数的定义. 这里补充一下.

定义 7.1. 设随机向量 $(X_1,\cdots,X_n)$ 的联合分布函数为 $F(x_1,\cdots,x_n)$, 用 ∞ 替换 $F(x_1,\cdots,x_n)$ 的一部分变量 (不是所有的变量) 得到的函数称为 $(X_1,\cdots,X_n)$ 的**边缘分布函数**.

联合分布函数唯一确定了所有的边缘分布函数, 而即使所有的边缘分布函数都已知, 联合分布函数仍然不能唯一确定.

例 7.1. 容易验证, 表 1 中的两个不同的分布具有相同的边缘分布函数.

p_{ij}	y_1	y_2
x_1	1/4	1/4
x_2	1/4	1/4

p_{ij}	y_1	y_2
x_1	1/3	1/6
x_2	1/6	1/3

表 1: 两个不同的分布

7.2 随机变量的独立性

补充几个定理的证明.

定理 7.1. 随机向量 $(X_1,\cdots,X_n)$ 相互独立当且仅当 $\forall x_1,x_2,\cdots,x_n$, 事件 $\{X_1\le x_1\},\{X_2\le x_2\},\cdots,\{X_n\le x_n\}$ 相互独立.

证明. 由书上的定义 1.2, 随机向量 $(X_1,\cdots,X_n)$ 相互独立当且仅当下式成立:

$$P(X_1\le x_1,X_2\le x_2,\cdots,X_n\le x_n)=P(X_1\le x_1)P(X_2\le x_2)\cdots P(X_n\le x_n). \tag{7.1}$$

($\Leftarrow$) 在书上第 2 章的定义 2.2 中取 $k=n$ 即得式 (7.1).

($\Rightarrow$) $\forall 1\le j_1<j_2<\cdots<j_k\le n$. 在式 (7.1) 中令 $\forall i\in\{1,2,\cdots,n\}\backslash\{j_1,j_2,\cdots,j_k\},x_i=\infty$, 得

$$P(X_{j_1}\le x_{j_1},X_{j_2}\le x_{j_2},\cdots,X_{j_k}\le x_{j_k})=P(X_{j_1}\le x_{j_1})P(X_{j_2}\le x_{j_2})\cdots P(X_{j_k}\le x_{j_k}).$$

由书上第 2 章的定义 2.2 得事件 $\{X_1\le x_1\},\{X_2\le x_2\},\cdots,\{X_n\le x_n\}$ 相互独立. □

7.3 连续型随机向量

如果联合密度是连续的, 那么所有的边缘密度都是连续的, 反之不成立.

例 7.2. 设 X 有连续的概率密度 $f(x)=\lambda e^{-\lambda x}$ $(\lambda>0)$, $Y=X$, 则随机向量只在射线 $l:y=x$ $(x>0)$ 上有非负取值. $\therefore$ 对于任意的 $\mathbb{R}^2$ 的长方形子集 D 和函数 f, 有

$$\iint_D f(x,y)\mathrm{d}x\mathrm{d}y=\iint_{l\cap D} f(x,y)\mathrm{d}x\mathrm{d}y.$$

$\because$ 直线是零测集, $\therefore \iint_{l\cap D} f(x,y)\mathrm{d}x\mathrm{d}y=0$.

考虑 $D=\{(x,y)|0<x\le 1,0<y\le 1\}$, 则 $P((X,Y)\in D)=P(0<X\le 1)=1-e^{-\lambda}>0$, 但 $\nexists f$ 使得 $\iint_D f(x,y)\mathrm{d}x\mathrm{d}y=1-e^{-\lambda}$. 由书上的定义 3.1, $\therefore (X,Y)$ 不是连续型随机向量, 没有联合密度.

类比离散型随机向量来理解书上的定理 3.2.

例 7.3. 设 X,Y 是取值为 $\mathbb{Z}$ 的离散型随机变量, $p_{ij}=P(X=i,Y=j)$, 有

$$\begin{aligned}p_{ij}&=P(X=i,Y\le j)-P(X=i,Y\le j-1)\\&=P(X\le i,Y\le j)-P(X\le i-1,Y\le j)-(P(X\le i,Y\le j-1)-P(X\le i-1,Y\le j-1)),\end{aligned}$$

可以将 $P(X=i,Y\le j)=P(X\le i,Y\le j)-P(X\le i-1,Y\le j)$ 理解为离散的 "偏导数", 将 p_{ij} 理解为离散的 "二阶混合偏导".

例 7.4 需要用到下面的定理.

定理 7.2 (书上定理 3.2 的推广)**.** 设 D 是 $\mathbb{R}^n$ 上的开集, $X_1,\cdots,X_n$ 的联合分布函数 $F(x_1,\cdots,x_n)$ 在 D 中可微, 且有连续的 n 阶混合偏导数. 如果 $P((X_1,\cdots,X_n)\in D)=1$, 那么

$$f(x_1,x_2,\cdots,x_n)=\begin{cases}\dfrac{\partial F(x_1,x_2,\cdots,x_n)}{\partial x_1\partial x_2\cdots\partial x_n}, & (x_1,x_2,\cdots,x_n)\in D,\\ 0 & \text{otherwise}.\end{cases}$$

证明. 略. □

可以把书上的例 3.2 推广到一般情形.

例 7.4. 一部手机陆续收到短信. 假设在不相交的时间段内收到的短信数相互独立, 且在任何长为 h 的时间段内收到的短信服从参数为 μh 的 Poisson 分布. 从 $t=0$ 开始, 用 X_i 表示第 i 个短信的到达时刻, 则 $X_1,X_2,\cdots,X_n$ 的概率密度为

$$f(x_1,x_2,\cdots,x_n)=\mu^n e^{-\mu x_n},\quad x_n>x_{n-1}>\cdots>x_1>0.$$

证明. 用数学归纳法. 由书上的例 3.2 得 $n=2$ 时成立. 假设 $X_1,X_2,\cdots,X_{n-1}$ 的概率密度为

$$g(x_1,x_2,\cdots,x_{n-1})=\mu^{n-1}e^{-\mu x_{n-1}},\quad x_{n-1}>\cdots>x_1>0.$$

对于 $x_n>x_{n-1}>\cdots>x_1>0$, 定义 $x_0=0$, 用 N_i 表示 $[x_{i-1},x_i]$ 内收到的短信数, 则 $N_1,N_2,\cdots,N_n$ 相互独立, 且 $N_i\sim\mathcal{P}(\mu(x_i-x_{i-1}))$. 有

$$\begin{aligned}P(X_1\le x_1,\cdots,X_{n-1}\le x_{n-1},X_n>x_n)&=P(N_1=1,\cdots,N_{n-1}=1,N_n=0)\\&=P(N_1=1)\cdots P(N_{n-1}=1)P(N_n=0)\\&=e^{-\mu(x_n-x_{n-1})}\prod_{i=1}^{n-1}\mu(x_i-x_{i-1})e^{-\mu(x_i-x_{i-1})}\\&=\mu^{n-1}e^{-\mu x_n}\prod_{i=1}^{n-1}(x_i-x_{i-1}).\end{aligned}$$

$\therefore$

$$\begin{aligned}F(x_1,x_2,\cdots,x_n)&=P(X_1\le x_1,\cdots,X_{n-1}\le x_{n-1})-P(X_1\le x_1,\cdots,X_{n-1}\le x_{n-1},X_n>x_n)\\&=P(X_1\le x_1,\cdots,X_{n-1}\le x_{n-1})-\mu^{n-1}e^{-\mu x_n}\prod_{i=1}^{n-1}(x_i-x_{i-1}).\end{aligned}$$

由定理 7.2, 当 $x_{n-1}>\cdots>x_1>0$ 时有

$$\begin{aligned}\frac{\partial F(x_1,x_2,\cdots,x_n)}{\partial x_1\partial x_2\cdots\partial x_{n-1}}&=g(x_1,x_2,\cdots,x_{n-1})-\frac{\partial\left(\mu^{n-1}e^{-\mu x_n}\prod\limits_{i=1}^{n-1}(x_i-x_{i-1})\right)}{\partial x_1\partial x_2\cdots\partial x_{n-1}}\\&=\mu^{n-1}e^{-\mu x_{n-1}}-\mu^{n-1}e^{-\mu x_n}\frac{\partial\left(\prod\limits_{i=1}^{n-1}(x_i-x_{i-1})\right)}{\partial x_1\partial x_2\cdots\partial x_{n-1}}.\end{aligned}$$

$\because$

$$\begin{aligned}\frac{\partial\left(\prod\limits_{i=1}^{n-1}(x_i-x_{i-1})\right)}{\partial x_1\partial x_2\cdots\partial x_{n-1}}&=\frac{\partial}{\partial x_1\partial x_2\cdots\partial x_{n-2}}\left(\frac{\partial}{\partial x_{n-1}}\left(\prod_{i=1}^{n-1}(x_i-x_{i-1})\right)\right)\\&=\frac{\partial}{\partial x_1\partial x_2\cdots\partial x_{n-2}}\left(\prod_{i=1}^{n-2}(x_i-x_{i-1})\cdot\frac{\partial(x_{n-1}-x_{n-2})}{\partial x_{n-1}}\right)\\&=\frac{\partial}{\partial x_1\partial x_2\cdots\partial x_{n-2}}\left(\prod_{i=1}^{n-2}(x_i-x_{i-1})\right)\\&=\cdots\\&=\frac{\partial(x_1-x_0)}{\partial x_1}=1,\end{aligned}$$

$\therefore$

$$\frac{\partial F(x_1,x_2,\cdots,x_n)}{\partial x_1\partial x_2\cdots\partial x_{n-1}}=\mu^{n-1}e^{-\mu x_{n-1}}-\mu^{n-1}e^{-\mu x_n},\quad x_{n-1}>\cdots>x_1>0.$$

由定理 7.2,

$$f(x_1,x_2,\cdots,x_n)=\frac{\partial}{\partial x_n}\left(\frac{\partial F(x_1,x_2,\cdots,x_n)}{\partial x_1\partial x_2\cdots\partial x_{n-1}}\right)=\mu^ne^{-\mu x_n},\quad x_n>x_{n-1}>\cdots>x_1>0.$$

□

7.4 连续型随机向量的独立性

设 X 有概率密度 $f_X(x)$. 考察那些使得 $P(X=x)>0$ 的 x 组成的集合. 由书上第 3 章的定义 2.2 得 $P(X=x)>0$ 当且仅当 $f_X(x)>0$. 称 $\{x|f_X(x)>0\}$ 为 X 的**支集**或**取值**. 如果随机向量 $(X_1,X_2,\cdots,X_n)$ 有概率密度 $f(x_1,x_2,\cdots,x_n)$, 则称 $\{(x_1,x_2,\cdots,x_n)|f(x_1,x_2,\cdots,x_n)>0\}$ 为 $(X_1,X_2,\cdots,X_n)$ 的支集.

对 (X,Y), 已知 $X=x_0$ 时, Y 的支集为 $\{y|f(x_0,y)>0\}$. 如果 X,Y 相互独立, 那么 Y 的支集为 $\{y|f_X(x_0)f_Y(y)>0\}$.

$\because X=x_0$, $\therefore f_X(x_0)>0$, $\therefore Y$ 的支集为 $\{y|f_Y(y)>0\}$. 取逆否命题就得到书上的定理 3.4 (1).

8 第 4 章习题

4.4. 常数 a 是随机变量, 按定义证明 a 与任何随机变量 Y 独立.

证明. $\because$ 当 $b \geq a$ 时有 $P(a \leq b) = 1, P(a \leq b, Y \leq y) = P(Y \leq y)$, 当 $b < a$ 时有 $P(a \leq b) = P(a \leq b, Y \leq y) = 0$, $\therefore \forall b, y$ 都有

$$P(a \leq b, Y \leq y) = P(Y \leq y)P(a \leq b).$$

由书上的定义 1.1 得 a 与 Y 独立. □

4.5. 证明: 对于固定的 x, 联合分布函数 $F(x,y)$ 关于 y 单调不减且右连续.

证明. 对 $y_1 \leq y_2$, 有 $\{X = x, Y = y_1\} \subset \{X = x, Y = y_2\}$. $\therefore F(x, y_1) < F(x, y_2)$. $\therefore F(x,y)$ 关于 y 单调不减.

$\because$ 关于 y 的函数 $F(x, y+\delta)$ 关于 δ 单调有界, $\therefore$ 当 $\delta \downarrow 0$ 时有极限, 且该极限等于任意递减趋于 y 的数列的函数值的极限. $\therefore$

$$\lim_{\delta \downarrow 0} F(x, y+\delta) = \lim_{n\to\infty} F\left(x, y + \frac{1}{n}\right) = \lim_{n\to\infty} P\left(\{X \leq x\} \cap \left\{Y \leq y + \frac{1}{n}\right\}\right)$$

由概率的连续性得

$$\begin{aligned}\lim_{n\to\infty} P\left(\{X \leq x\} \cap \left\{Y \leq y + \frac{1}{n}\right\}\right) &= P\left(\bigcap_{n=1}^{\infty}\left(\{X \leq x\} \cap \left\{Y \leq y + \frac{1}{n}\right\}\right)\right) \\ &= P\left(\{X \leq x\} \cap \left(\bigcap_{n=1}^{\infty}\left\{Y \leq y + \frac{1}{n}\right\}\right)\right) \\ &= P(\{X \leq x\} \cap \{Y \leq y\}) = F(x,y).\end{aligned}$$

$\therefore F(x,y)$ 关于 y 右连续. □

4.6. 设 $X \sim \mathcal{B}(n,p), Y \sim \mathcal{B}(m,p)$, X, Y 相互独立, 计算 $X+Y$ 的概率分布.

解. 对 $0 \leq z \leq m+n$, 有

$$P(X+Y=z) = \sum_{i+j=z} P(X=i, Y=j).$$

$\because X, Y$ 相互独立, $\therefore$

$$\begin{aligned}\sum_{i+j=z} P(X=i, Y=j) &= \sum_{i+j=z} P(X=i)P(Y=j) \\ &= \sum_{i+j=z} C_n^i p^i (1-p)^{n-i} C_m^j p^j (1-p)^{m-j} \\ &= p^z (1-p)^{m+n-z} \sum_{i+j=z} C_n^i C_m^j.\end{aligned}$$

考察等式

$$(1+x)^n(1+x)^m = (1+x)^{m+n}$$

两边的展开式, 左边的 x^k 一项的系数为 $\sum\limits_{i+j=z} C_n^i C_m^j$, 右边的 x^k 一项的系数为 C_{m+n}^k, $\therefore$

$$\sum_{i+j=z} C_n^i C_m^j = C_{m+n}^k.$$

$\therefore$

$$\sum_{i+j=z} P(X=i, Y=j) = C_{m+n}^k p^z (1-p)^{m+n-z},$$

$\therefore X+Y \sim \mathcal{B}(m+n, p)$.

4.9. 设随机变量 X, Y 独立同分布, 证明

$$P(a < \min(X,Y) \le b) = P^2(X>a) - P^2(X>b).$$

证明.

$$\begin{aligned} P(a < \min(X,Y) \le b) &= P(\min(X,Y) > a) - P(\min(X,Y) > b) \\ &= P(X>a, Y>a) - P(X>b, Y>b). \end{aligned}$$

$\because X, Y$ 相互独立, $\therefore$

$$P(X>a, Y>a) - P(X>b, Y>b) = P(X>a)P(Y>a) - P(X>b)P(Y>b).$$

$\because X, Y$ 同分布, $\therefore P(Y>x) = P(X>x)$,

$$P(X>a)P(Y>a) - P(X>b)P(Y>b) = P^2(X>a) - P^2(X>b).$$ □

4.11. 设 a 是常数, (X,Y) 有概率密度

$$f(x,y) = \begin{cases} ax^2y, & x^2 < y < 1, \\ 0, & \text{otherwise.} \end{cases}$$

求 X, Y 的边缘密度, 说明 X, Y 不独立.

解. 由

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\mathrm{d}x\mathrm{d}y = \iint\limits_{x^2<y<1} ax^2y\mathrm{d}x\mathrm{d}y = 1$$

得 $a = \dfrac{21}{4}$. 有

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\mathrm{d}y = \int_{x^2}^{1} \frac{21}{4}x^2y\mathrm{d}y = \frac{21}{8}x^2(1-x^2), \quad -1<x<1,$$

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\mathrm{d}x = \int_{-\sqrt{y}}^{\sqrt{y}} \frac{21}{4}x^2y\mathrm{d}x = \frac{7}{2}y^2\sqrt{y}, \quad 0<y<1.$$

$\because f_X(x)f_Y(y) \ne f(x,y)$, $\therefore X, Y$ 不独立.

4.14. 设随机变量 X, Y 相互独立, $X \sim \mathcal{E}(\lambda), Y \sim \mathcal{E}(\mu)$, 计算 $P(X > Y)$.

解. $\because X, Y$ 相互独立且分别有概率密度, 由书上的定理 3.1 得 (X, Y) 有概率密度

$$f(x,y) = f_X(x) f_Y(y) = \lambda e^{-\lambda x} \mu e^{-\mu y} = \lambda\mu e^{-(\lambda x + \mu y)}, \quad x > 0, y > 0.$$

$\therefore$

$$\begin{aligned} P(X > Y) &= \iint_{\mathbb{R}_+^2} f(x,y) I[x > y] \mathrm{d}x\mathrm{d}y \\ &= \int_0^\infty \mathrm{d}x \int_0^x f(x,y)\mathrm{d}y \\ &= \int_0^\infty \mathrm{d}x \int_0^x \lambda\mu e^{-(\lambda x + \mu y)}\mathrm{d}y \\ &= \int_0^\infty \lambda e^{-\lambda x}(1 - e^{-\mu x})\mathrm{d}x \\ &= 1 - \frac{\lambda}{\lambda + \mu} = \frac{\mu}{\lambda + \mu}. \end{aligned}$$

4.16. 设 $p = 1 - q \in (0, 1)$, $0 < \alpha < q/p$. 假设一个家庭有 n 个小孩的概率为

$$p_n = \begin{cases} \alpha p^n, & n \geq 1, \\ 1 - \alpha p/q, & n = 0. \end{cases}$$

如果男婴和女婴的出生是等可能的且相互独立, 求一个家庭有 n 个男孩的概率.

解. 设 X 为一个家庭中的小孩的个数, Y 为一个家庭中的男孩的个数, 则 $P(X = x) = p_x$.

由全概率公式得

$$P(Y = y) = \sum_{x=0}^\infty P(Y = y | X = x) P(X = x).$$

$\because$ 男婴和女婴的出生是等可能的且相互独立, $\therefore$

$$P(Y = y | X = x) = \begin{cases} C_x^y \left(\frac{1}{2}\right)^y \left(\frac{1}{2}\right)^{x-y}, & x \geq y, \\ 0, & \text{otherwise} \end{cases} = \begin{cases} C_x^y \frac{1}{2^x}, & x \geq y, \\ 0, & \text{otherwise.} \end{cases}$$

$\therefore$

$$
\begin{aligned}
P(Y=0) &= \left(1-\alpha\frac{p}{q}\right)C_0^0\frac{1}{2^0}+\sum_{x=0}^{\infty}C_x^y\frac{1}{2^x}\cdot\alpha p^x \\
&= 1-\alpha\frac{p}{q}+\sum_{x=1}^{\infty}\alpha C_x^0\left(\frac{p}{2}\right)^x \\
&= 1-\alpha\frac{p}{q}+\alpha\sum_{x=1}^{\infty}\left(\frac{p}{2}\right)^x \\
&= 1-\alpha\frac{p}{q}+\alpha\frac{p/2}{1-p/2} \\
&= 1-\alpha\frac{p}{q}+\alpha\frac{p}{1+q},
\end{aligned}
$$

当 $y\neq 0$ 时有

$$
P(Y=y)=\sum_{x=y}^{\infty}C_x^y\frac{1}{2^x}\cdot\alpha p^x
$$

4.17. 设随机向量 (X,Y) 有联合密度

$$
f(x,y)=\begin{cases}\dfrac{1+xy}{4}, & x,y\in(-1,1),\\ 0 & \text{otherwise.}\end{cases}
$$

证明 X^2,Y^2 相互独立, 但 X,Y 不独立.

解. $\because\ \forall x_0\in(-1,1)$,

$$
\frac{f(x,y)}{f(x_0,y)}=\frac{1+xy}{1+x_0y}
$$

都是 y 的函数, 由书上的定理 3.4 (2) 得 X,Y 不独立.

令 $U=X^2,V=Y^2$, 则 $X=\sqrt{U},Y=\sqrt{V}$,

$$
\frac{\partial(u,v)}{\partial(x,y)}=\begin{vmatrix}2x & 0\\ 0 & 2y\end{vmatrix}=4xy,\quad \frac{\partial(x,y)}{\partial(u,v)}=\left(\frac{\partial(u,v)}{\partial(x,y)}\right)^{-1}=\frac{1}{4xy}=\frac{1}{4\sqrt{uv}}.
$$

$\therefore$

$$\begin{aligned}P(U=u,V=v) &= P(X=\sqrt{u},Y=\sqrt{v})+P(X=-\sqrt{u},Y=\sqrt{v})\\&\quad+P(X=\sqrt{u},Y=-\sqrt{v})+P(X=-\sqrt{u},Y=-\sqrt{v})\\&=\frac{1+\sqrt{u}\sqrt{v}}{4}\mathrm{d}x\mathrm{d}y+\frac{1+(-\sqrt{u})\sqrt{v}}{4}\mathrm{d}x\mathrm{d}y\\&\quad+\frac{1+(-\sqrt{u})(-\sqrt{v})}{4}\mathrm{d}x\mathrm{d}y+\frac{1+\sqrt{u}(-\sqrt{v})}{4}\mathrm{d}x\mathrm{d}y\\&=\frac{4+2\sqrt{u}\sqrt{v}-2\sqrt{u}\sqrt{v}}{4}\cdot\frac{\partial(x,y)}{\partial(u,v)}\mathrm{d}u\mathrm{d}v\\&=\frac{1}{4\sqrt{uv}}\mathrm{d}u\mathrm{d}v.\end{aligned}$$

$\therefore (U,V)$ 有联合密度

$$g(u,v)=\frac{1}{4\sqrt{uv}},u,v\in(0,1).$$

U 的边缘密度为

$$f_U(u)=\int_0^1\frac{1}{4\sqrt{uv}}\mathrm{d}v=\frac{1}{2\sqrt{u}}.$$

对称地, $f_V(v)=\dfrac{1}{2\sqrt{v}}$. $\because g(u,v)=f_U(u)f_V(v)$, $\therefore U,V$ 相互独立.

4.18. 设 D 是非负连续函数 $g(x)$ 与 x 轴所夹的区域, D 的面积 $m(D)\in(0,\infty)$. 设 (X,Y) 在 D 上均匀分布, 求 X 的概率密度.

解. 设 $x_0\in\mathbb{R}$. 有

$$\begin{aligned}P(X\le x_0)&=\iint_D\frac{1}{m(D)}I[X\le x_0]\mathrm{d}x\mathrm{d}y\\&=\int_{-\infty}^{x_0}\mathrm{d}x\int_0^{g(x)}\frac{1}{m(D)}\mathrm{d}y\\&=\int_{-\infty}^{x_0}\frac{g(x)}{m(D)}\mathrm{d}x.\end{aligned}$$

$\therefore X$ 有概率密度

$$f_X(x)=\frac{g(x)}{m(D)},$$

其中 $m(D)=\int_{-\infty}^{\infty}g(x)\mathrm{d}x$.

4.19. 设随机变量 X 服从二项分布 $\mathcal{B}(n,p)$, Y 服从指数分布 $\mathcal{E}(\lambda)$. 当 X,Y 独立时, 求 $Z=Y-X$ 的分布函数和概率密度.

解. $\because Y\sim\mathcal{E}(\lambda),\therefore P(Y\le y)=(1-e^{-\lambda y})I[y\ge 0]$.

由全概率公式得

$$\begin{aligned}P(Y-X\le z)&=\sum_{k=0}^{n}P(Y-X\le z|X=k)P(X=k)\\&=\sum_{k=0}^{n}P(Y\le z+k|X=k)P(X=k).\end{aligned}$$

$\because X,Y$ 相互独立, $\therefore$

$$\begin{aligned}\sum_{k=0}^{n}P(Y\le z+k|X=k)P(X=k)&=\sum_{k=0}^{n}P(Y\le z+k)P(X=k)\\&=\sum_{k=0}^{n}(1-e^{-\lambda(z+k)})I[z+k\ge 0]C_n^kp^k(1-p)^{n-k}\\&=\sum_{k=0}^{n}(1-e^{-\lambda(z+k)})I[z\ge -k]C_n^kp^k(1-p)^{n-k}.\end{aligned}$$

$\therefore$

$$\begin{aligned}F_Z(z)&=\sum_{k=0}^{n}(1-e^{-\lambda(z+k)})I[z\ge -k]C_n^kp^k(1-p)^{n-k},\\f_Z(z)&=\frac{\mathrm{d}F_Z(z)}{\mathrm{d}z}=\lambda e^{-\lambda(z+k)}I[z\ge -k]C_n^kp^k(1-p)^{n-k}.\end{aligned}$$

4.22. 设随机变量 X,Y 独立, X 有概率密度 $f(x)$, Y 有离散型概率分布 $P(Y=a_i)=p_i>0$ $(i=1,2,\cdots)$. 证明:

(1) 若 $a_1,a_2,\cdots$ 都不为 0, 证明 $Z=XY$ 有概率密度

$$h(z)=\sum_{i=1}^{\infty}\frac{p_i}{|a_i|}f\left(\frac{z}{a_i}\right).$$

(2) 若有某个 $a_i=0$, 则 Z 不是连续型随机变量.

证明. (1) 由全概率公式得

$$\begin{aligned}P(Z=z)&=\sum_{i=1}^{\infty}P(XY=z|Y=a_i)P(Y=a_i)\\&=\sum_{i=1}^{\infty}P\left(X=\frac{z}{a_i}\Big|Y=a_i\right)P(Y=a_i).\end{aligned}$$

$\because X,Y$ 独立, $\therefore$

$$\begin{aligned}\sum_{i=1}^{\infty}P\left(X=\frac{z}{a_i}\Big|Y=a_i\right)P(Y=a_i)&=\sum_{i=1}^{\infty}P\left(X=\frac{z}{a_i}\right)P(Y=a_i)\\&=\sum_{i=1}^{\infty}f\left(\frac{z}{a_i}\right)\cdot\frac{1}{|a_i|}p_i.\end{aligned}$$

(2) 不妨设 $a_1=0$. $\because \{Z=0\}=\{XY=0\}\supset\{Y=0\}\supset\{Y=a_1\}$, $\therefore P(Z=0)\ge P(Y=a_1)>0$. $\therefore Z$ 不是连续型随机变量. □

4.23. 设随机向量 (X,Y) 有联合密度

$$f(x,y)=\begin{cases}\frac{1}{2}(x+y)e^{-(x+y)}, & x,y>0\\ 0 & \text{otherwise.}\end{cases}$$

求 $Z=X+Y$ 的概率密度.

解. 令 $W=X$, 则 $X=W, Y=Z-W$,

$$\frac{\partial(z,w)}{\partial(x,y)}=\begin{vmatrix}1 & 1\\ 1 & 0\end{vmatrix}=-1,\quad \frac{\partial(x,y)}{\partial(z,w)}=\left(\frac{\partial(z,w)}{\partial(x,y)}\right)^{-1}=-1.$$

$$P(Z=z,W=w)=P(X=w,Y=z-w)=f(w,z-w)\mathrm{d}x\mathrm{d}y=\frac{1}{2}ze^{-z}\mathrm{d}w\mathrm{d}z,\quad w>0,z>0,$$

$\therefore$

$$f_Z(z)=\frac{1}{2}ze^{-z},\quad z>0.$$

4.25. 设随机变量 U,V 独立, 都服从 $(0,1)$ 上的均匀分布.

(a) 计算 $R=\sqrt{-2\ln U}$ 和 $\Theta=2\pi V$ 的概率密度;

(b) 证明 $X=R\cos\Theta, Y=R\sin\Theta$ 独立, 都服从 Gauss 分布.

证明. (a) R 的取值为 $(0,\infty)$, Θ 的取值为 $(0,2\pi)$. 有

$$P(R=r)=P(\sqrt{-2\ln U}=r)=P(U=e^{-r^2/2})=|-re^{-r^2/2}|=re^{-r^2/2},\quad r\in(0,\infty),$$

$$P(\Theta=\theta)=P\left(V=\frac{\theta}{2\pi}\right)=\frac{1}{2\pi},\quad \theta\in(0,2\pi).$$

(b) (X,Y) 的取值为 $\{X^2+Y^2\neq 0\}$. 由 $X=R\cos\Theta, Y=R\sin\Theta$ 得 $R=\sqrt{X^2+Y^2}$,

$$\Theta=g(X,Y)=\begin{cases}0 & Y=0,X>0\\ \pi & Y=0,X<0\\ \arctan\frac{Y}{X} & Y\neq 0,X\geq 0\\ \pi-\arctan\frac{Y}{X} & Y\neq 0,X<0\end{cases}.$$

有

$$\frac{\partial(r,\theta)}{\partial(x,y)}=\left(\frac{\partial(x,y)}{\partial(r,\theta)}\right)^{-1}=\frac{1}{r}.$$

$\therefore$

$$\begin{aligned}P(X=x,Y=y)&=P(R=\sqrt{x^2+y^2},\Theta=g(X,Y))\\&=\frac{1}{2\pi}re^{-r^2/2}\mathrm{d}r\mathrm{d}\theta\\&=\frac{1}{2\pi}re^{-r^2/2}\cdot\frac{1}{r}\mathrm{d}x\mathrm{d}y\\&=\frac{1}{2\pi}e^{-(x^2+y^2)/2}\mathrm{d}x\mathrm{d}y.\end{aligned}$$

$\therefore (X,Y)$ 有概率密度

$$f(x,y)=\frac{1}{2\pi}e^{-(x^2+y^2)/2}=\frac{1}{\sqrt{2\pi}}e^{-x^2/2}\cdot\frac{1}{\sqrt{2\pi}}e^{-y^2/2}.$$

容易验证 X,Y 都服从 Gauss 分布且相互独立. □

证明第 4.27 题需要先证明下面的引理.

引理 8.1. 设取值为 $\mathbb{R}$ 的连续型随机变量 X,Y,Z 相互独立, 则 $\forall a,b\in\mathbb{R}$, $aX+bY$ 与 Z 相互独立.[a]

[a]这可以由书上的定理 1.1 (3) 得到, 由于那里没有证明, 所以还是单独拿出来作为一个引理.

证明. $\because X,Y,Z$ 是连续型随机变量, $\therefore (X,Y,Z)$ 有概率密度 $f(x,y,z)$. $\because X,Y,Z$ 相互独立, $\therefore$

$$f(x,y,z)=f_X(x)f_Y(y)f_Z(z)=f_{X,Y}(x,y)f_Z(z),$$

其中 $f_{X,Y}(x,y)$ 是 (X,Y) 的联合密度. 设

$$\begin{pmatrix}U\\V\end{pmatrix}=\boldsymbol{A}\begin{pmatrix}X\\Y\end{pmatrix},$$

其中

$$\boldsymbol{A}=\begin{pmatrix}a&b\\c&d\end{pmatrix},\quad \boldsymbol{A}^{-1}=\begin{pmatrix}a_1&b_1\\c_1&d_1\end{pmatrix},$$

与第 4.33 题类似, 有

$$\begin{aligned}P(U=u,V=v,Z=z)&=P(X=a_1u+b_1v,Y=c_1u+d_1v,Z=z)\\&=f_{X,Y}(a_1u+b_1v,c_1u+d_1v)f_Z(z)\mathrm{d}u\mathrm{d}v\mathrm{d}z\\&=f_{U,V}(u,v)f_Z(z)\mathrm{d}u\mathrm{d}v\mathrm{d}z,\end{aligned}$$

其中 $f_{U,V}(u,v)$ $(u,v\in\mathbb{R})$ 是 (U,V) 的联合密度. $\therefore$

$$\begin{aligned}P(U\le u_0,Z\le z_0)&=P(U\le u_0,V\le\infty,Z\le z_0)\\&=\int_{-\infty}^{z_0}f_Z(z)\mathrm{d}z\int_{-\infty}^{u_0}\mathrm{d}u\int_{-\infty}^{\infty}f_{U,V}(u,v)\mathrm{d}v\\&=\int_{-\infty}^{z_0}f_Z(z)\mathrm{d}z\int_{-\infty}^{u_0}f_U(u)\mathrm{d}u\\&=P(U\le u_0)P(Z\le z_0).\end{aligned}$$

$\therefore U=aX+bY$ 与 Z 相互独立. □

4.27. 设随机变量 $X_1, X_2, \cdots, X_n$ 相互独立, 都服从 Gauss 分布, $a_1, a_2, \cdots, a_n$ 是不全为 0 的常数, 证明 $U = a_1X_1 + a_2X_2 + \cdots + a_nX_n$ 服从 Gauss 分布.

证明. 用数学归纳法. $n = 1$ 时的情形是显然的.

假设对相互独立的随机变量 $X_1, \cdots, X_{n-1}$ 和任意不全为 0 的常数 $a_1, \cdots, a_{n-1}$, $U' = a_1X_1 + \cdots + a_{n-1}X_{n-1}$ 服从 Gauss 分布.

$\because X_1, X_2, \cdots, X_n$ 相互独立, 由引理 8.1 得 $a_1X_1 + a_2X_2, X_3, \cdots, X_n$ 相互独立, $\therefore 1 \cdot (a_1X_1 + a_2X_2) + a_3X_3, X_4, \cdots, X_n$ 相互独立, $\cdots$, $\therefore U'$ 与 X_n 相互独立.

由书上的定理 6.2 (3) 得 (U', X) 服从二维 Gauss 分布, 由书上的定理 6.2 (5) 得 $U = 1 \cdot U' + a_nX_n$ 服从 Gauss 分布. □

4.28 (b)**.** 设随机向量 (X_1, X_2) 服从二维 Gauss 分布, 有联合密度

$$f(x_1, x_2) = \frac{\exp\left(-\frac{1}{2(1-\rho^2)}\left(\frac{(x_1-\mu_1)^2}{\sigma_1^2} - \frac{2\rho(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\right)\right)}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}},$$

设 $f_1(x)$ 是 X_1 的边缘密度, 验证

$$\frac{f(x_1, x_2)}{f_1(x_1)} = \frac{1}{\sqrt{2\pi(1-\rho^2)}\sigma_2} \exp\left(-\frac{(x_2-\mu_x)^2}{2(1-\rho^2)\sigma_2^2}\right)$$

关于 x_2 是 Gauss 概率密度, 其中 $\mu_x = \mu_2 + (\rho\sigma_2/\sigma_1)(x_1 - \mu_1)$.

证明. 由书上的定理 6.2 (1) 得

$$f_1(x_1) = \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left(-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right).$$

$\therefore$

$$\frac{f(x_1, x_2)}{f_1(x_1)} = \frac{1}{\sqrt{2\pi(1-\rho^2)}\sigma_2} e^u,$$

其中

$$\begin{aligned}
u &= -\frac{1}{2(1-\rho^2)}\left(\frac{(x_1-\mu_1)^2}{\sigma_1^2} - \frac{2\rho(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\right) + \frac{(x-\mu_1)^2}{2\sigma_1^2} \\
&= -\frac{1}{2(1-\rho^2)}\left((1-(1-\rho^2))\frac{(x_1-\mu_1)^2}{\sigma_1^2} - \frac{2\rho(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\right) \\
&= -\frac{1}{2(1-\rho^2)}\left(\frac{\rho^2(x_1-\mu_1)^2}{\sigma_1^2} - \frac{2\rho(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\right) \\
&= -\frac{1}{2(1-\rho^2)}\left(\frac{\rho(x_1-\mu_1)}{\sigma_1} - \frac{x_2-\mu_2}{\sigma_2}\right)^2 \\
&= -\frac{1}{2(1-\rho^2)\sigma_2^2}\left(x_2-\mu_2-\frac{\rho\sigma_2(x_1-\mu_1)}{\sigma_1}\right)^2 \\
&= -\frac{1}{2(1-\rho^2)\sigma_2^2}(x_2-\mu_x)^2.
\end{aligned}$$

□

4.30. 设随机变量 X, Y 相互独立, $X \sim \mathcal{E}(\lambda), Y \sim \mathcal{E}(\mu)$. 求 $U = \min(X, Y)$, $V = \max(X, Y)$ 和 $W = X/Y$ 的概率密度.

解. 有

$$\begin{aligned} P(\min(X,Y) \le u) &= P(\{X \le u\} \cup \{Y \le u\}) \\ &= P(X \le u) + P(Y \le u) - P(X \le u, Y \le u). \end{aligned}$$

$\because X, Y$ 相互独立, $\therefore$

$$\begin{aligned} P(X \le u) + P(Y \le u) - P(X \le u, Y \le u) &= P(X \le u) + P(Y \le u) - P(X \le u)P(Y \le u) \\ &= 1 - (P(X \le u) - 1)(P(Y \le u) - 1) \\ &= 1 - (1 - e^{-\lambda u} - 1)(1 - e^{-\mu u} - 1) \\ &= 1 - e^{-(\lambda+\mu)u}. \end{aligned}$$

$\therefore U$ 有概率密度

$$f_U(u) = (\lambda + \mu)e^{-(\lambda+\mu)u}, \quad u \in (0, \infty).$$

有

$$\begin{aligned} P(\max(X,Y) \le v) &= P(X \le v, Y \le v) \\ &= P(X \le v)P(Y \le v) \\ &= (1 - e^{-\lambda v})(1 - e^{-\mu v}) \\ &= 1 - e^{-\lambda v} - e^{-\mu v} + e^{-(\lambda+\mu)v}. \end{aligned}$$

$\therefore V$ 有概率密度

$$f_V(v) = \lambda e^{-\lambda v} + \mu e^{-\mu v} - (\lambda + \mu)e^{-(\lambda+\mu)v}, \quad v \in (0, \infty).$$

$\because X, Y$ 相互独立, $\therefore (X, Y)$ 有联合密度

$$f(x, y) = \lambda e^{-\lambda x} \mu e^{-\mu y} = \lambda\mu e^{-(\lambda x + \mu y)}.$$

设 $Z = Y$, 则 $X = WZ$, $\dfrac{\partial(x,y)}{\partial(w,z)} = z$. 对 $w, z \in (0, \infty)$, 有

$$\begin{aligned} P(Z = z, W = w) &= P(X = wz, Y = z) \\ &= f(wz, z)\mathrm{d}x\mathrm{d}y \\ &= \lambda\mu e^{-(\lambda w + \mu)z} z\mathrm{d}w\mathrm{d}z. \end{aligned}$$

$\therefore$

$$P(W \le w_0) = P(Z \le \infty, W \le w_0) = \int_0^{w_0} \mathrm{d}w \int_0^{\infty} \lambda\mu e^{-(\lambda w + \mu)z} z\mathrm{d}z.$$

$\therefore$

$$
\begin{aligned}
f_W(w) &= \int_0^\infty \lambda\mu e^{-(\lambda w+\mu)z} z\mathrm{d}z \\
&= -\frac{\lambda\mu}{\lambda w+\mu}\int_0^\infty z\mathrm{d}e^{-(\lambda w+\mu)z} \\
&= \frac{\lambda\mu}{\lambda w+\mu}\int_0^\infty e^{-(\lambda w+\mu)z}\mathrm{d}z \\
&= \frac{\lambda\mu}{(\lambda w+\mu)^2}, \quad w\in(0,\infty).
\end{aligned}
$$

注. 在平面上作出 $\min(X,Y)\le u$ 和 $\max(X,Y)\le v$ 对应的区域如图 4 中的蓝色区域. 可以对区域 D_1, D_2 积分得到 $\min(X,Y), \max(X,Y)$ 的分布函数.

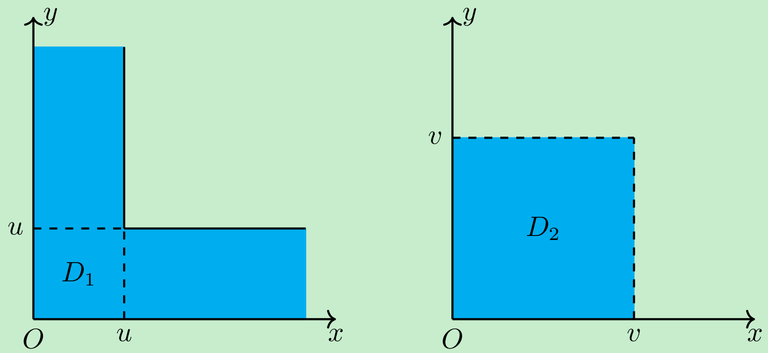

图 4: 左图为 $D_1=\{\min(X,Y)\le u\}$, 右图为 $D_2=\{\max(X,Y)\le v\}$

4.31. 设随机向量 (X,Y) 有联合密度 $f(x)g(y)$, (U,V) 有联合密度

$$
p(u,v)=\begin{cases}\dfrac{1}{\alpha}f(u)g(v), & u\ge v,\\ 0 & u<v,\end{cases}
$$

(a) 求 U,V 的边缘密度;

(b) 证明 $\alpha=P(X\ge Y)$.

证明. (a) $\because$

$$
\begin{aligned}
P(U\le u) &= \iint\limits_{x\le u} \frac{1}{\alpha}f(x)g(y)I[x<y]\mathrm{d}x\mathrm{d}y \\
&= \int_{-\infty}^u \frac{1}{\alpha}f(x)\mathrm{d}x\int_{-\infty}^x g(y)\mathrm{d}y \\
&= \int_{-\infty}^u \frac{1}{\alpha}f(x)G(x)\mathrm{d}x,
\end{aligned}
$$

其中 $G(x)=\int_{-\infty}^{x} g(y)\mathrm{d}y$, $\therefore U$ 的概率密度为 $f_U(u)=\dfrac{1}{\alpha}f(x)G(x)$. $\because$

$$\begin{aligned}
P(V\le v)&=1-P(V>v)\\
&=1-\iint\limits_{y>v}\frac{1}{\alpha}f(x)g(y)I[x<y]\mathrm{d}x\mathrm{d}y\\
&=1-\int_v^{\infty}\frac{1}{\alpha}g(y)\mathrm{d}y\int_y^{\infty}f(x)\mathrm{d}x\\
&=1-\int_v^{\infty}\frac{1}{\alpha}g(y)F(y)\mathrm{d}y,
\end{aligned}$$

其中 $F(y)=\int_{-\infty}^{y} f(x)\mathrm{d}x$, $\therefore U$ 的概率密度为 $f_V(v)=\dfrac{1}{\alpha}g(y)F(y)$.

(b) 有

$$\begin{aligned}
P(X\ge Y)&=\iint\limits_{x>y}f(x)g(y)\mathrm{d}x\mathrm{d}y\\
&=\alpha\int_{-\infty}^{\infty}\frac{1}{\alpha}f(x)\mathrm{d}x\int_{-\infty}^{x}g(y)\mathrm{d}y\\
&=\alpha\cdot P(U\le\infty)=\alpha.
\end{aligned}$$

□

4.32. 随机变量 X,Y 相互独立, X 有分布函数 $F_X(x)$ 和概率密度 $f_X(x)=F_X'(x)$. 如果 $P(Y>y)=(P(X>y))^{\beta}$, 其中 β 是正常数, 求 $P(X\ge Y)$.

解. Y 的分布函数为

$$F_Y(y)=1-P(Y>y)=1-(P(X>y))^{\beta}=1-(1-F_X(x))^{\beta}.$$

$\because X$ 有概率密度 $f_X(x)$, $\therefore F_X(x)$ 可导. $\therefore F_Y(y)$ 可导. $\therefore Y$ 有概率密度 $f_Y(y)=F_Y'(y)$. 有

$$\begin{aligned}
P(X\ge Y)&=\iint\limits_{x\ge y}f_X(x)f_Y(y)\mathrm{d}x\mathrm{d}y\\
&=\int_{-\infty}^{\infty}f_X(x)\mathrm{d}x\int_{-\infty}^{x}f_Y(y)\mathrm{d}y\\
&=\int_{-\infty}^{\infty}f_X(x)F_Y(x)\mathrm{d}x\\
&=\int_{-\infty}^{\infty}F_X'(x)(1-(1-F_X(x))^{\beta})\mathrm{d}x.
\end{aligned}$$

令 $t=F_X(x)$, 则 $x=\infty$ 时 $t=1$, $x=-\infty$ 时 $t=0$. 有

$$\int_{-\infty}^{\infty}F_X'(x)(1-(1-F_X(x))^{\beta})\mathrm{d}x=\int_0^1(1-(1-t)^{\beta})\mathrm{d}t=\frac{\beta}{1+\beta}.$$

4.33 (有修改)**.** 证明: 如果 $\mathbb{R}^2$ 上的随机向量 (X,Y) 有联合密度 $f(x,y)$, (U,V) 由线性变换

$$\begin{pmatrix} U \\ V \end{pmatrix} = \boldsymbol{A} \begin{pmatrix} X \\ Y \end{pmatrix} + \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$$

决定, 其中

$$\boldsymbol{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad \boldsymbol{A}^{-1} = \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix},$$

则 (U,V) 有联合密度

$$g(u,v) = \frac{1}{|\det \boldsymbol{A}|} f((u-\mu_1, v-\mu_2)\boldsymbol{A}^{-T}), \quad (u,v) \in \mathbb{R}^2,$$

其中 $\boldsymbol{A}^{-T} = (A^{-1})^T$.

证明. $\forall u, v$, 有

$$\begin{aligned} P(U=u, V=v) &= P\left(\boldsymbol{A}\begin{pmatrix} X \\ Y \end{pmatrix} + \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} = \begin{pmatrix} u \\ v \end{pmatrix}\right) \\ &= P\left(\begin{pmatrix} X \\ Y \end{pmatrix} = \boldsymbol{A}^{-1}\begin{pmatrix} u-\mu_1 \\ v-\mu_2 \end{pmatrix}\right). \end{aligned}$$

$\because \dfrac{\partial(x,y)}{\partial(u,v)} = |\boldsymbol{A}^{-1}|$, $\therefore$

$$\begin{aligned} P\left(\begin{pmatrix} X \\ Y \end{pmatrix} = \boldsymbol{A}^{-1}\begin{pmatrix} u-\mu_1 \\ v-\mu_2 \end{pmatrix}\right) &= f\left(\left(\boldsymbol{A}^{-1}\begin{pmatrix} u-\mu_1 \\ v-\mu_2 \end{pmatrix}\right)^T\right)|\det \boldsymbol{A}^{-1}|\mathrm{d}u\mathrm{d}v \\ &= f((u-\mu_1, v-\mu_2)\boldsymbol{A}^{-T})|\det \boldsymbol{A}^{-1}|\mathrm{d}u\mathrm{d}v \\ &= \frac{1}{|\det \boldsymbol{A}|} f((u-\mu_1, v-\mu_2)\boldsymbol{A}^{-T})\mathrm{d}u\mathrm{d}v. \end{aligned}$$

$\therefore$

$$g(u,v) = \frac{1}{|\det \boldsymbol{A}|} f((u-\mu_1, v-\mu_2)\boldsymbol{A}^{-T}). \qquad \square$$

9 第 5 章笔记

可以证明, 任意的概率空间 $(\Omega, \mathcal{F}, P)$ 上所有有二阶矩的随机变量构成的线性空间 (其中的加法和纯量乘法是随机变量的加法和纯量乘法) 是一个 Hilbert 空间, 内积为 $\langle X, Y \rangle = E(XY)$. 这个 Hilbert 空间是无穷维的. 第 7 章第 6 节介绍的均方收敛可以看成是在这个 Hilbert 空间中收敛.

书上的例 4.5 用到了下面的引理.

引理 9.1. 设 X 是随机变量, 如果 $\int_0^\infty P(|X| > x)\mathrm{d}x = 0$, 那么 $\forall x > 0, P(|X| > x) = 0$.

证明. 假设 $\exists x_0>0$ 使得 $P(|X|>x_0)=\varepsilon>0$, 则 $\forall x\in(0,x_0), P(|X|>x)\ge P(|X|>x_0)=\varepsilon$. $\therefore$

$$\int_0^{\infty}P(|X|>x)\mathrm{d}x\ge\int_0^{x_0}P(|X|>x)\mathrm{d}x\ge x_0\varepsilon>0.$$

这与 $\int_0^{\infty}P(|X|>x)\mathrm{d}x=0$ 矛盾. □

补充几个定理的证明.

定理 9.1 (书上的定理 5.2)**.** 设 a,b,c 是常数, $\mu_j=EX_j, \mathrm{Var}(X_j)<\infty$ $(1\le j\le n)$, 则

(4) $\mathrm{Var}\left(\sum\limits_{j=1}^n X_j\right)=\sum\limits_{i=1}^n\sum\limits_{j=1}^n(E(X_iX_j)-\mu_i\mu_j)$;

(5) 当 $X_1,X_2,\cdots,X_n$ 相互独立时, $\mathrm{Var}\left(\sum\limits_{j=1}^n X_j\right)=\sum\limits_{j=1}^n\mathrm{Var}(X_j)$.

证明. (4) 有

$$\begin{aligned}
\mathrm{Var}\left(\sum_{j=1}^n X_j\right)&=E\left(\sum_{j=1}^n X_j-\sum_{j=1}^n\mu_j\right)^2\\
&=E\left(\left(\sum_{j=1}^n X_j\right)^2+\left(\sum_{j=1}^n\mu_j\right)^2-2\left(\sum_{j=1}^n X_j\right)\left(\sum_{j=1}^n\mu_j\right)\right)\\
&=E\left(\left(\sum_{j=1}^n X_j\right)^2-\left(\sum_{j=1}^n\mu_j\right)^2-2\left(\sum_{j=1}^n X_j-\sum_{j=1}^n\mu_j\right)\left(\sum_{j=1}^n\mu_j\right)\right)\\
&=E\left(\sum_{j=1}^n X_j\right)^2-\left(\sum_{j=1}^n\mu_j\right)^2-E\left(2\left(\sum_{j=1}^n X_j-\sum_{j=1}^n\mu_j\right)\left(\sum_{j=1}^n\mu_j\right)\right)\\
&=E\left(\sum_{i=1}^n\sum_{j=1}^n X_iX_j\right)-\left(\sum_{i=1}^n\sum_{j=1}^n\mu_i\mu_j\right)-2\left(\sum_{j=1}^n\mu_j\right)\left(\sum_{j=1}^n EX_j-\sum_{j=1}^n\mu_j\right)\\
&=\left(\sum_{i=1}^n\sum_{j=1}^n E(X_iX_j)\right)-\left(\sum_{i=1}^n\sum_{j=1}^n\mu_i\mu_j\right)-2\left(\sum_{j=1}^n\mu_j\right)\sum_{j=1}^n(EX_j-\mu_j)\\
&=\sum_{i=1}^n\sum_{j=1}^n(E(X_iX_j)-\mu_i\mu_j).
\end{aligned}$$

(5) $\because X_1,X_2,\cdots,X_n$ 相互独立, $\therefore\forall i,j=1,2,\cdots,n, j\ne i, E(X_iX_j)=EX_iEX_j$. 由 (4) 得

$$\begin{aligned}
\mathrm{Var}\left(\sum_{j=1}^n X_j\right)&=\sum_{i=1}^n\sum_{j=1}^n(E(X_iX_j)-\mu_i\mu_j)\\
&=\sum_{i=1}^n\sum_{\substack{j=1\\ j\ne i}}^n(EX_iEX_j-\mu_i\mu_j)+\sum_{i=1}^n(EX_i^2-\mu_i^2)\\
&=\sum_{i=1}^n(EX_i^2-(EX_i)^2)=\sum_{i=1}^n\mathrm{Var}(X_i).
\end{aligned}$$

□

定理 9.2 (书上的定理 6.2)**.** 设 ρ_{XY} 是 X,Y 的相关系数, 则有

(1) $|\rho_{XY}| \le 1$;

(2) $|\rho_{XY}| = 1$ 的充要条件为有常数 a,b 使得 $P(Y = a + bX) = 1$;

(3) 如果 X,Y 独立, 则 X,Y 不相关.

证明. (1) 由定义,

$$|\rho_{XY}| = \frac{|E((X-\mu_X)(Y-\mu_Y))|}{\sigma_X \sigma_Y}.$$

由内积不等式 (书上的定理 6.1),

$$|E((X-\mu_X)(Y-\mu_Y))| \le \sqrt{E(X-\mu_X)^2 E(Y-\mu_Y)^2} = \sigma_X\sigma_Y, \tag{9.1}$$

$\therefore |\rho_{XY}| \le 1$.

(2) 由内积不等式取等号的充要条件得式 (9.1) 成立的充要条件是有常数 a,b 使得 $P(Y - \mu_Y = a + b(X - \mu_X)) = 1$, 即有常数 $b, a' = a + \mu_Y - b\mu_X$ 使得 $P(Y = a' + bX)$.

(3) $\because X,Y$ 独立, $\therefore$ 由书上的定理 4.1 得 $E(XY) = EXEY$. 由书上的式 (6.6) 得

$$\sigma_{XY} = E(XY) - EXEY = 0. \qquad \square$$

设随机向量 (X,Y) 服从二维 Gauss 分布, 则分布由 X 的均值 μ_X, 方差 σ_X, Y 的均值 μ_Y, 方差 σ_Y, X,Y 的相关系数 ρ_{XY} 这 5 个量完全确定. $\therefore$ 可以用 $(X,Y) \sim N(\mu_X,\mu_Y;\sigma_X,\sigma_Y;\rho_{XY})$ 来表示 (X,Y) 服从由这 5 个量确定的二维 Gauss 分布.

10 第 5 章习题

5.5. 一部手机收到的短信中有 $p = 2\%$ 是广告, 你期望相邻的两次广告短信中有多少个不是广告短信?

解. 设 X 为相邻的两次广告短信中不是广告短信的数目, 则 $P(X=k)=p^k(1-p)$. 有

$$\begin{aligned}EX&=\sum_{k=0}^{\infty}kP(X=k)\\&=\sum_{k=1}^{\infty}kp^k(1-p)\\&=p(1-p)\sum_{k=1}^{\infty}kp^{k-1}\\&=p(1-p)\left(\sum_{k=1}^{\infty}\frac{\partial}{\partial x}x^k\right)\Bigg|_{x=p}\\&=p(1-p)\left(\frac{\partial}{\partial x}\sum_{k=0}^{\infty}x^k\right)\Bigg|_{x=p}\\&=p(1-p)\left(\frac{\partial}{\partial x}\frac{1}{1-x}\right)\Bigg|_{x=p}\\&=p(1-p)\cdot\frac{1}{(1-p)^2}=\frac{p}{1-p}.\end{aligned}$$

把 $p=2\%$ 代入得 $EX=49$.

5.8. 设 X,Y 独立, 都服从 Gauss 分布, 求 $R=\sqrt{X^2+Y^2}$ 的数学期望.

解. 由书上的定理 4.3 得 R 有 Rayleigh 概率密度

$$f_R(r)=re^{-r^2/2}.$$

$\therefore$

$$\begin{aligned}ER&=\int_{-\infty}^{\infty}r^2e^{-r^2/2}\mathrm{d}r\\&=\int_{-\infty}^{\infty}re^{-r^2/2}\mathrm{d}r^2/2\\&=\sqrt{2}\int_{0}^{\infty}t^{1/2}e^{-t}\mathrm{d}t\\&=\sqrt{2}\Gamma(3/2)=\sqrt{\frac{\pi}{2}}.\end{aligned}$$

5.9. 设 $X_1,X_2,\cdots,X_n$ 独立同分布, 且 $P(X_i>0)=1$. 对于 $k\le n$, 计算

$$E\frac{X_1+X_2+\cdots+X_k}{X_1+X_2+\cdots+X_n}.$$

解. 设

$$Y_i=\frac{X_i}{X_1+X_2+\cdots+X_n},\quad (n=1,2,\cdots,n).$$

$\because X_1, X_2, \cdots, X_n$ 独立同分布, $\therefore Y_1, Y_2, \cdots, Y_n$ 同分布. $\because P(X_i > 0) = 1, \therefore P(Y_i > 0) = 1$. 有

$$1 = E\frac{X_1 + X_2 + \cdots + X_n}{X_1 + X_2 + \cdots + X_n} = EY_1 + EY_2 + \cdots + EY_n = nEY_1,$$

$$E\frac{X_1 + X_2 + \cdots + X_k}{X_1 + X_2 + \cdots + X_n} = EY_1 + EY_2 + \cdots + EY_k = kEY_1,$$

$\therefore$

$$E\frac{X_1 + X_2 + \cdots + X_k}{X_1 + X_2 + \cdots + X_n} = \frac{k}{n}.$$

5.12. 设 $X_1, X_2, \cdots, X_n$ 相互独立, 有共同的离散分布 $p_k = P(X = k)$, $k = 1, 2, \cdots$. 引入 $u_k = p_1 + p_2 + \cdots + p_{k-1}, v_k = 1 - u_k$, 证明:

$$E[\min(X_1, X_2, \cdots, X_n)] = \sum_{k=1}^{\infty} v_k^n, \quad E[\max(X_1, X_2, \cdots, X_n)] = \sum_{k=1}^{\infty} (1 - u_k^n).$$

证明. 由书上的定理 3.2 (3) 得

$$\begin{aligned}
E[\min(X_1, X_2, \cdots, X_n)] &= \sum_{k=1}^{\infty} P(\min(X_1, X_2, \cdots, X_n) \geq k) \\
&= \sum_{k=1}^{\infty} P(X_1 \geq k, X_2 \geq k, \cdots, X_n \geq k) \\
&= \sum_{k=1}^{\infty} P(X_1 \geq k) P(X_2 \geq k) \cdots P(X_n \geq k) \\
&= \sum_{k=1}^{\infty} (1 - P(X_1 \geq k))(1 - P(X_2 \geq k)) \cdots (1 - P(X_n \geq k)) \\
&= \sum_{k=1}^{\infty} v_k^n,
\end{aligned}$$

$$\begin{aligned}
E[\max(X_1, X_2, \cdots, X_n)] &= \sum_{k=1}^{\infty} P(\max(X_1, X_2, \cdots, X_n) \geq k) \\
&= \sum_{k=1}^{\infty} 1 - P(\max(X_1, X_2, \cdots, X_n) < k) \\
&= \sum_{k=1}^{\infty} 1 - P(X_1 < k, X_2 < k, \cdots, X_n < k) \\
&= \sum_{k=1}^{\infty} 1 - P(X_1 < k) P(X_2 < k) \cdots P(X_n < k) \\
&= \sum_{k=1}^{\infty} 1 - u_k^n.
\end{aligned}$$

□

5.14 (1). 设 (X,Y) 有概率密度

$$f(x,y)=\begin{cases}\dfrac{3}{2x^2y^3}, & x>1, 1<xy<x^2,\\ 0, & \text{otherwise.}\end{cases}$$

计算 EY.

解. 有 $1<xy<x^2 \Rightarrow \dfrac{1}{x}<y<x$. 由书上的定理 3.1 (2),

$$\begin{aligned}EY &= \iint_{\mathbb{R}^2} yf(x,y)\mathrm{d}x\mathrm{d}y\\ &= \int_1^{\infty}\mathrm{d}x\int_{1/x}^{x} y\frac{3}{2x^3y^2}\mathrm{d}y\\ &= \int_1^{\infty}\frac{3}{2x^3}\mathrm{d}x\int_{1/x}^{x}\frac{1}{y}\mathrm{d}y\\ &= \int_1^{\infty}\frac{3(\ln x-\ln(1/x))}{2x^3}\mathrm{d}x\\ &= \int_1^{\infty}\frac{3\ln x}{x^2}\mathrm{d}\ln x\\ &= \int_0^{\infty}3te^{-2t}\mathrm{d}t=\frac{3}{4}.\end{aligned}$$

5.17 (b). 设 5 台计算机独立工作, 每台计算机感染病毒前的时间 (分别记作 $X_1,\cdots,X_5$) 都服从参数为 λ 的指数分布 $\mathcal{E}(\lambda)$. 对 5 台计算机都感染病毒前的时间的数学期望是多少?

解. 记 $Y=\max(X_1,\cdots,X_5)$, 则

$$\begin{aligned}P(Y>y) &= 1-P(Y\le y)\\ &= 1-P(X_1\le y, X_2\le y,\cdots,X_5\le y)\\ &= 1-P(X_1\le y)P(X_2\le y)\cdots P(X_5\le y)\\ &= 1-(1-e^{-\lambda y})^5\\ &= 1-\left(1+\sum_{i=1}^{5}C_5^i(-e^{-\lambda y})^i\right)\\ &= \sum_{i=1}^{5}C_5^i(-1)^{i-1}e^{-\lambda iy}.\end{aligned}$$

由书上的定理 3.1 (3) 得

$$\begin{aligned}
EY &= \int_0^\infty P(Y>y)\mathrm{d}y \\
&= \int_0^\infty \sum_{i=1}^5 C_5^i(-1)^{i-1}e^{-\lambda iy}\mathrm{d}y \\
&= \sum_{i=1}^5 C_5^i(-1)^{i-1}\int_0^\infty e^{-\lambda iy}\mathrm{d}y \\
&= \sum_{i=1}^5 C_5^i\frac{1}{\lambda i}(-1)^{i-1} = \frac{137}{60\lambda}.
\end{aligned}$$

5.20. 设 $n\ge 2$, $X_1, X_2, \cdots, X_n$ 独立同分布,

$$\hat{\mu} = \frac{1}{n}\sum_{j=1}^n X_j, \quad \hat{\sigma}^2 = \frac{1}{n-1}\sum_{j=1}^n (X_j-\hat{\mu})^2,$$

验证 $E\hat{\mu} = EX_1, E\hat{\sigma}^2 = \mathrm{Var}(X_1)$.

证明. $\because X_1, X_2, \cdots, X_n$ 独立同分布, $\therefore$

$$E\hat{\mu} = \frac{1}{n}\sum_{j=1}^n EX_j = \frac{1}{n}\sum_{j=1}^n EX_1 = EX_1.$$

$\therefore$

$$\begin{aligned}
E\hat{\sigma}^2 &= \frac{1}{n-1}\sum_{j=1}^n E(X_j-\hat{\mu})^2 \\
&= \frac{1}{n-1}\sum_{j=1}^n E\left(X_j - \frac{1}{n}\sum_{i=1}^n X_i\right)^2 \\
&= \frac{1}{n-1}\sum_{j=1}^n E\left(\frac{n-1}{n}X_j - \frac{1}{n}\sum_{\substack{i=1\\ i\ne j}}^n X_i\right)^2 \\
&= \frac{1}{n-1}\sum_{j=1}^n E\left(\frac{n-1}{n}(X_j-EX_1) - \frac{1}{n}\sum_{\substack{i=1\\ i\ne j}}^n (X_i-EX_1)\right)^2.
\end{aligned}$$

令 $Y_i = X_i - EX_1$ $(i=1,2,\cdots,n)$, 则 Y_i 独立同分布, 均值为 0, 方差为 $\mathrm{Var}(X_1)$. $\therefore \forall i =$

$1,\cdots,n, j=1,\cdots,\hat{i},\cdots,n, E(X_iX_j)=E(X_i)E(X_j)=0$. 有

$$
\begin{aligned}
E\hat{\sigma}^2 &= \frac{1}{n-1}\sum_{j=1}^{n} E\left(\frac{n-1}{n}Y_j - \frac{1}{n}\sum_{\substack{i=1\\ i\neq j}}^{n} Y_i\right)^2 \\
&= \frac{1}{n-1}\sum_{j=1}^{n} E\left(\left(\frac{n-1}{n}Y_j\right)^2 + \left(\frac{1}{n}\sum_{\substack{i=1\\ i\neq j}}^{n} Y_i\right)^2 - 2\frac{n-1}{n}Y_j\left(\frac{1}{n}\sum_{\substack{i=1\\ i\neq j}}^{n} Y_i\right)\right) \\
&= \frac{1}{n-1}\sum_{j=1}^{n}\left(\left(\frac{n-1}{n}\right)^2 EY_j^2 + \frac{1}{n^2}\sum_{\substack{i=1\\ i\neq j}}^{n}\sum_{\substack{k=1\\ k\neq j}}^{n} E(Y_iY_k) - 2\frac{n-1}{n^2}\sum_{\substack{i=1\\ i\neq j}}^{n} E(Y_iY_j)\right) \\
&= \frac{1}{n-1}\sum_{j=1}^{n}\left(\left(\frac{n-1}{n}\right)^2 EY_j^2 + \frac{1}{n^2}\sum_{\substack{i=1\\ i\neq j}}^{n} EY_i^2 + \frac{1}{n^2}\sum_{\substack{i=1\\ i\neq j}}^{n}\sum_{\substack{k=1\\ k\neq j\\ k\neq i}}^{n} E(Y_iY_k) - 2\frac{n-1}{n^2}\sum_{\substack{i=1\\ i\neq j}}^{n} E(Y_iY_j)\right) \\
&= \frac{1}{n-1}\sum_{j=1}^{n}\left(\left(\frac{n-1}{n}\right)^2 E(Y_j-EY_j)^2 + \frac{1}{n^2}\sum_{\substack{i=1\\ i\neq j}}^{n} E(Y_i-EY_i)^2\right) \\
&= \frac{1}{n-1}\sum_{j=1}^{n}\left(\left(\frac{n-1}{n}\right)^2 \mathrm{Var}(X_1) + \frac{1}{n^2}\sum_{\substack{i=1\\ i\neq j}}^{n} \mathrm{Var}(X_1)\right) \\
&= \frac{1}{n-1}\sum_{j=1}^{n}\left(\frac{(n-1)^2}{n^2}\mathrm{Var}(X_1) + \frac{n-1}{n^2}\mathrm{Var}(X_1)\right) \\
&= \sum_{j=1}^{n}\frac{1}{n}\mathrm{Var}(X_1) = \mathrm{Var}(X_1).
\end{aligned}
$$

□

5.22. 设 $X_1, X_2, \cdots, X_n$ 是相互独立的随机变量, $\mathrm{Var}(X_i)=\sigma_i^2$. 求满足 $\sum\limits_{j=1}^{n} a_j = 1, a_j \geq 0$ 的常数 $a_1, a_2, \cdots, a_n$ 使得 $Y=\sum\limits_{j=1}^{n} a_jX_j$ 的方差最小.

解. $\because X_1, X_2, \cdots, X_n$ 相互独立, $\therefore$

$$
\begin{aligned}
P(a_1X_1\leq x_1, a_2X_2\leq x_2,\cdots,a_nX_n\leq x_n) &= P(X_1\leq x_1/a_1, X_2\leq x_2/a_2,\cdots,X_n\leq x_n/a_n) \\
&= P(X_1\leq x_1/a_1)P(X_2\leq x_2/a_2)\cdots P(X_n\leq x_n/a_n) \\
&= P(a_1X_1\leq x_1)P(a_2X_2\leq x_2)\cdots P(a_nX_n\leq x_n).
\end{aligned}
$$

$\therefore a_1X_1, a_2X_2, \cdots, a_nX_n$ 相互独立. 由定理 9.1 (5) 得

$$\mathrm{Var}(Y)=\sum_{j=1}^{n}\mathrm{Var}(a_jX_j)=\sum_{j=1}^{n} a_j^2\sigma_i^2.$$

由 Cauchy 不等式得

$$1=\sum_{j=1}^{n}a_j\sigma_j\cdot\frac{1}{\sigma_j}\le\left(\sum_{j=1}^{n}(a_j\sigma_i)^2\right)\left(\sum_{j=1}^{n}\frac{1}{\sigma_j^2}\right),$$

当且仅当 $a_j\sigma_j=\lambda\cdot\dfrac{1}{\sigma_j}$, $j=1,2,\cdots,n$ 时取等号. $\therefore a_j=\dfrac{\sigma_j^{-2}}{\sum_{i=1}^{n}\sigma_i^{-2}}$, $j=1,2,\cdots,n$.

5.25. 设非负随机变量 X 有离散分布 $p_j=P(X=x_j),j\ge 1$. 证明:

(1) $P(X>x)=\sum_{j=1}^{\infty}p_jI[x<x_j]$;

(2) $EX=\int_0^{\infty}P(X>x)\mathrm{d}x$.

证明. (1) 有 $\forall j,P(X>x|X=x_j)=I[x<x_j]$. 由全概率公式得

$$P(X>x)=\sum_{j=1}^{\infty}P(X>x|X=x_j)P(X=x_j)=\sum_{j=1}^{\infty}p_jI[x<x_j].$$

(2) 由 (1) 得

$$\int_0^{\infty}P(X>x)\mathrm{d}x=\int_0^{\infty}\sum_{j=1}^{\infty}p_jI[x<x_j]\mathrm{d}x.$$

$\because \sum_{j=1}^{\infty}p_jI[x<x_j]\le\sum_{j=1}^{\infty}p_j=1$, 由 Weierstrass 判别法得 $\sum_{j=1}^{\infty}p_jI[x<x_j]$ 在 $(0,\infty)$ 上一致收敛, $\therefore$ 求和号和积分号可以交换次序. 有

$$\begin{aligned}\int_0^{\infty}\sum_{j=1}^{\infty}p_jI[x<x_j]\mathrm{d}x&=\sum_{j=1}^{\infty}p_j\int_0^{\infty}I[x<x_j]\mathrm{d}x\\&=\sum_{j=1}^{\infty}p_j\int_0^{x_j}\mathrm{d}x\\&=\sum_{j=1}^{\infty}p_jx_j=EX.\end{aligned}$$

□

5.26. 如果 $\mathrm{Cov}(X,Y)=0$, 验证 $\rho(X+Y,X-Y)=\dfrac{\mathrm{Var}(X)-\mathrm{Var}(Y)}{\mathrm{Var}(X)+\mathrm{Var}(Y)}$.

证明. 定义 $E^2X:=(EX)^2$. 由书上的式 (6.6) 得

$$\begin{aligned}\mathrm{Cov}(X+Y,X-Y)&=E((X+Y)(X-Y))-E(X+Y)E(X-Y)\\&=E(X^2-Y^2)-(EX+EY)(EX-EY)\\&=EX^2-EY^2-(E^2X-E^2Y)\\&=(EX^2-E^2X)-(EY^2-E^2Y)\\&=\mathrm{Var}(X)-\mathrm{Var}(Y).\end{aligned}$$

由书上的式 (5.2) 得

$$\begin{aligned}\operatorname{Var}(X \pm Y) &= E(X \pm Y)^2 - E^2(X \pm Y) \\ &= E(X^2 + Y^2 \pm 2XY) - (EX \pm EY)^2 \\ &= EX^2 + EY^2 \pm 2E(XY) - (E^2X + E^2Y \pm 2EXEY) \\ &= (EX^2 - E^2X) + (EY^2 - E^2Y) \pm 2(E(XY) - EXEY) \\ &= \operatorname{Var}(X) + \operatorname{Var}(Y) \pm 2\operatorname{Cov}(X, Y).\end{aligned}$$

$\because \operatorname{Cov}(X, Y) = 0$, $\therefore \operatorname{Var}(X \pm Y) = \operatorname{Var}(X) + \operatorname{Var}(Y)$,

$$\begin{aligned}\rho(X + Y, X - Y) &= \frac{\operatorname{Cov}(X + Y, X - Y)}{\sqrt{\operatorname{Var}(X + Y)\operatorname{Var}(X - Y)}} \\ &= \frac{\operatorname{Var}(X) - \operatorname{Var}(Y)}{\sqrt{(\operatorname{Var}(X) + \operatorname{Var}(Y))^2}} \\ &= \frac{\operatorname{Var}(X) - \operatorname{Var}(Y)}{\operatorname{Var}(X) + \operatorname{Var}(Y)}.\end{aligned}$$

□

5.28. 设 (X, Y) 有概率密度

$$f(x, y) = \begin{cases} x + y & x \in (0, 1), y \in (0, 1), \\ 0 & \text{otherwise.} \end{cases}$$

计算 $\operatorname{Cov}(X, Y)$.

解. 有

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\mathrm{d}y = \int_0^1 (x + y)\mathrm{d}y = x + \frac{1}{2}, \quad x \in (0, 1),$$

$$EX = \int_{-\infty}^{\infty} x f_X(x)\mathrm{d}x = \frac{7}{12}.$$

对称地, 有 $EY = \dfrac{7}{12}$.

$$E(XY) = \iint_{\mathbb{R}^2} xyf(x, y)\mathrm{d}x\mathrm{d}y = \int_0^1 \mathrm{d}x \int_0^1 xy(x + y)\mathrm{d}y = \frac{1}{3}.$$

$\therefore$

$$\operatorname{Cov}(X, Y) = E(XY) - EXEY = -\frac{1}{144}.$$

5.30. 设活塞的直径 X 的平均值为 20.00cm, 标准差为 0.02; 气缸的直径 Y 的平均值为 20.10cm, 标准差为 0.02. 设 X 与 Y 独立且都服从 Gauss 分布, 计算活塞能装入气缸的概率.

解. 事件“活塞能装入气缸”等价于事件 $A=\{X<Y\}$. 设 (X,Y) 的联合密度为 $f(x,y)$, 则

$$P(A)=\iint_{\mathbb{R}^2} f(x,y)I[x<y]\mathrm{d}x\mathrm{d}y=\iint_{x<y} f(x,y)\mathrm{d}x\mathrm{d}y.$$

由题得

$$f(x,y)=\frac{1}{2\pi\cdot 0.02^2}\exp\left(-\frac{(x-20.0)^2}{2\cdot 0.02^2}-\frac{(y-20.1)^2}{2\cdot 0.02^2}\right).$$

$\therefore$

$$P(A)=\int_{-\infty}^{\infty}\mathrm{d}y\int_{-\infty}^{y}\frac{1}{2\pi\cdot 0.02^2}\exp\left(-\frac{(x-20.0)^2}{2\cdot 0.02^2}-\frac{(y-20.1)^2}{2\cdot 0.02^2}\right)\mathrm{d}x\approx 0.9998.$$

注. 我不会算这个积分, 所以用了 Mathematica 的数值积分来计算. 理论上, 对 $\forall a$

$$\int_{-\infty}^{\infty}\mathrm{d}y\int_{-\infty}^{y}\frac{1}{2\pi\cdot 0.02^2}\exp\left(-\frac{(x-a)^2}{2\cdot 0.02^2}-\frac{(y-a-0.1)^2}{2\cdot 0.02^2}\right)\mathrm{d}x$$

的值都一样, 都为 $P(A)$, 但是在 Mathematica 中, 上式取不同的 a, 计算结果不一样.

取 $a=-10,0,10,20$ 分别计算上式的 Mathematica 程序如下:

```
a = {-10, 0, 10, 20};
NIntegrate[
  1/(2 Pi 0.02^2)*Exp[-(x - a)^2/(2*0.02^2) - (y - a - 0.1)^2/(2*0.02^2)],
  {y, -Infinity, Infinity}, {x, -Infinity, y},
  MinRecursion -> 100, MaxRecursion -> 100]
```

5.32. 设 X 的概率密度是偶函数, $0<EX^2<\infty$, 证明 $|X|$ 与 X 不相关也不独立.

证明. $\because X$ 的概率密度是偶函数, $\therefore EX=0$, $\therefore$

$$\mathrm{Cov}(X,|X|)=E(X|X|)-EXE|X|=E(X|X|).$$

设 X 的概率密度为 $f(x)$. $\because f(x)$ 是偶函数, $\therefore x|x|f(x)$ 是奇函数,

$$E(X|X|)=\int_{-\infty}^{\infty}x|x|f(x)\mathrm{d}x=0.$$

$\therefore X,|X|$ 不相关.

若 $X=x$, 则 $|X|$ 的取值只能为 $|x|$. $\therefore X,|X|$ 不独立. □

5.33. 设 $X\sim N(0,\sigma^2)$, $n\in\mathbb{N}_+$, 验证

$$EX^n=\begin{cases}\sigma^n(n-1)!!, & n=2m,\\ 0 & \text{otherwise}.\end{cases}$$

证明. 有

$$EX^n = \int_{-\infty}^{\infty} x^n \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2} \mathrm{d}x.$$

当 $n = 2m+1$ 时, 被积函数是奇函数, $\therefore EX^n = 0$. 当 $n = 2m$ 时, 被积函数是偶函数, 有

$$EX^n = 2\int_0^{\infty} x^n \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2} \mathrm{d}x.$$

令 $t = \dfrac{x^2}{2\sigma^2}$, 则 $x = \sigma\sqrt{2t}, \mathrm{d}x = \dfrac{\sigma}{\sqrt{2t}}\mathrm{d}t$. 代入得

$$\begin{aligned}
2\int_0^{\infty} x^n \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2} \mathrm{d}x &= 2\int_0^{\infty} (\sigma\sqrt{2t})^n \frac{1}{\sqrt{2\pi}\sigma} e^{-t} \frac{\sigma}{\sqrt{2t}} \mathrm{d}t \\
&= 2\frac{\sigma^n}{\sqrt{2\pi}} \int_0^{\infty} (\sqrt{2t})^{n-1} e^{-t} \mathrm{d}t \\
&= 2\frac{\sigma^n 2^{n/2-1}}{\sqrt{\pi}} \int_0^{\infty} t^{(n-1)/2} e^{-t} \mathrm{d}t \\
&= \frac{\sigma^n 2^{n/2}}{\sqrt{\pi}} \Gamma\left(\frac{n+1}{2}\right).
\end{aligned}$$

$\because \Gamma(\alpha+1) = \alpha\Gamma(\alpha)$, $\therefore$

$$\begin{aligned}
\Gamma\left(\frac{n+1}{2}\right) &= \frac{n-1}{2} \cdot \frac{n-3}{2} \cdots \frac{1}{2}\Gamma\left(\frac{1}{2}\right) \\
&= \frac{1}{2^{n/2}}(n-1)!! \cdot \sqrt{\pi}.
\end{aligned}$$

$\therefore EX = \sigma^n (n-1)!!$. □

5.35 (以及 5.18)**.** 设点随机地落在中心为原点, 半径为 R 的圆周上, 落点坐标是 (X, Y), 求 X, Y 的方差和协方差.

解. 由第 3.26 题得 x 有概率密度

$$f(x) = \frac{2\sqrt{R^2 - x^2}}{\pi R^2}, \quad -R < x < R.$$

$\because f(x)$ 是偶函数, $\therefore EX = \int_{-R}^{R} x f(x) \mathrm{d}x = 0$. 对称地, $EY = 0$. $\therefore \mathrm{Var}(X) = EX^2, \mathrm{Var}(Y) = EY^2, \mathrm{Cov}(X, Y) = E(XY)$.

由对称性得 $EX^2 = EY^2$, $\therefore$

$$EX^2 = \frac{EX^2 + EY^2}{2} = \frac{E(X^2 + Y^2)}{2}.$$

有

$$\begin{aligned}
E(X^2 + Y^2) &= \iint\limits_{x^2+y^2 \le R^2} (x^2 + y^2) \frac{1}{\pi R^2} \mathrm{d}x\mathrm{d}y \\
&= \frac{1}{\pi R^2} \int_0^{2\pi} \mathrm{d}\theta \int_0^R r^2 \cdot r\mathrm{d}r \\
&= \frac{R^2}{2}.
\end{aligned}$$

$\therefore EX^2 = EY^2 = \dfrac{R^2}{4}$.
有

$$E(XY) = \iint\limits_{x^2+y^2\le R^2} xy\frac{1}{\pi R^2}\mathrm{d}x\mathrm{d}y = 0.$$

注. 书上的答案应该有点问题, 我用 Mathematica 算 $\int_{-R}^{R} x^2 f(x)\mathrm{d}x$ 的结果也是 $\dfrac{R^2}{4}$.

第 5.36 题需要用到下面的引理.

引理 10.1. 设 $X_1, X_2, \cdots, X_n$ 相互独立, 都服从参数为 λ 的指数分布, 则 $Z = X_1 + X_2 + \cdots + X_n$ 有密度函数

$$f_Z(z) = \frac{\lambda^n}{(n-1)!}e^{-\lambda z}z^{n-1}, \quad z > 0.$$

证明. 用数学归纳法. $n = 1$ 时的情形由指数分布的定义得. 假设对符合题目要求的 $X_1, \cdots, X_{n-1}, X_n$, $W = X_1 + X_2 + \cdots + X_{n-1}$ 有密度函数

$$f_W(w) = \frac{\lambda^{n-1}}{(n-2)!}e^{-\lambda z}z^{n-2}, \quad z > 0.$$

令 $Z = W + X_n = X_1 + X_2 + \cdots + X_n$, 则

$$P(Z \le z) = \iint\limits_{\substack{x+w\le z\\x>0\\w>0}} P(X_n = x, W = w).$$

由引理 8.1 可以归纳地得到[2]: W 与 X_n 独立. $\therefore P(X_n = x, W = w) = P(X_n = x)P(W = w) = f_{X_n}(x)f_W(w)\mathrm{d}x\mathrm{d}w$, 其中 $f_{X_n}(x)$ 是 X_n 的概率密度. $\therefore$

$$\begin{aligned}
\iint\limits_{\substack{x+w\le z\\x>0\\w>0}} P(X_n = x, W = w) &= \iint\limits_{\substack{x+w\le z\\x>0\\w>0}} f_{X_n}(x)f_W(w)\mathrm{d}x\mathrm{d}w \\
&= \int_0^z f_W(w)\mathrm{d}w\int_0^{z-w} f_{X_n}(x)\mathrm{d}x \\
&= \int_0^z \frac{\lambda^{n-1}}{(n-2)!}e^{-\lambda w}w^{n-2}\mathrm{d}w\int_0^{z-w} \lambda e^{-\lambda x}\mathrm{d}x \\
&= \frac{\lambda^{n-1}}{(n-2)!}\int_0^z e^{-\lambda w}(1 - e^{-\lambda(z-w)})w^{n-2}\mathrm{d}w \\
&= \frac{\lambda^{n-1}}{(n-2)!}\int_0^z (e^{-\lambda w} - e^{-\lambda z})w^{n-2}\mathrm{d}w \\
&= \frac{\lambda^{n-1}}{(n-2)!}\left(\int_0^z e^{-\lambda w}w^{n-2}\mathrm{d}w - e^{-\lambda z}\int_0^z w^{n-2}\mathrm{d}w\right).
\end{aligned}$$

[2] 参考第 4.27 题.

$\therefore$

$$\begin{aligned}
f_Z(z) &= \frac{\partial}{\partial z} P(Z \le z) \\
&= \frac{\lambda^{n-1}}{(n-2)!} \cdot \frac{\partial}{\partial z} \left(\int_0^z e^{-\lambda w} w^{n-2} \mathrm{d}w - e^{-\lambda z} \int_0^z w^{n-2} \mathrm{d}w \right) \\
&= \frac{\lambda^{n-1}}{(n-2)!} \left(e^{-\lambda z} z^{n-2} - e^{-\lambda z} z^{n-2} - (-\lambda) e^{-\lambda z} \int_0^z w^{n-2} \mathrm{d}w \right) \\
&= \frac{\lambda^n}{(n-2)!} e^{-\lambda z} \int_0^z w^{n-2} \mathrm{d}w \\
&= \frac{\lambda^n}{(n-1)!} e^{-\lambda z} z^{n-1}.
\end{aligned}$$

$\therefore$ 原命题对 n 个变量也成立. □

5.36 (a)**.** 假设你的手机收到短信的时间间隔是相互独立的随机变量, 都服从参数为 $\lambda = 1/2$ 的指数分布. 现在你在等一个朋友的短信, 如果每个短信以概率 $p = 0.1$ 来自你的这位朋友, 从 $t = 0$ 开始, 用 Y 表示等待时间的长度. 计算等待时间 Y 的概率分布.

解. 设你朋友的短信是第 X 个到达的短信, 则 X 服从几何分布, $P(X = k) = pq^{k-1}$, $k = 1, 2, \cdots, q = 1 - p$. 由全概率公式得

$$P(Y = y) = \sum_{n=1}^{\infty} P(Y = y | X = n) P(X = n) = \sum_{n=1}^{\infty} P(Y = y | X = n) pq^{n-1}.$$

在你朋友的短信是第 n 个到达的短信的情况下, 等待时间是 n 个独立的指数分布之和. 由引理 10.1 得

$$\begin{aligned}
\sum_{n=1}^{\infty} P(Y = y | X = n) pq^{n-1} &= \sum_{n=1}^{\infty} \frac{\lambda^n}{(n-1)!} e^{-\lambda y} y^{n-1} pq^{n-1} \\
&= \sum_{n=1}^{\infty} \frac{\lambda^n}{(n-1)!} e^{-\lambda y} y^{n-1} pq^{n-1} \\
&= \lambda p e^{-\lambda y} \sum_{n=1}^{\infty} \frac{(\lambda y q)^{n-1}}{(n-1)!} \\
&= \lambda p e^{-\lambda y} \sum_{n=0}^{\infty} \frac{(\lambda y q)^n}{n!} \\
&= \lambda p e^{-\lambda y} e^{\lambda y q} \\
&= \lambda p e^{-\lambda p y}.
\end{aligned}$$

把 $\lambda = 1/2, p = 0.1$ 代入得 $Y \sim \mathcal{E}(1/20)$.

11 第 6 章笔记

由书上的定理 3.1 (3) 得

$$EX = E(E(X|Y)) = \int_{-\infty}^{\infty} E(X|Y = y) P(Y = y) \mathrm{d}y.$$

这可以看成是书上的定理 1.1 (1) 对连续型随机变量的推广.

条件数学期望 $E(X|Y)$ 按定义是用一个 $\Omega\to\mathbb{R}$ 的函数 Y 代替 $\mathbb{R}$ 上的函数 $m(y):=E(X|Y=y)$ 中的 y. 由于 $\Omega\to\mathbb{R}$ 的函数与 $\mathbb{R}$ 上的函数复合仍然是 $\Omega\to\mathbb{R}$ 的函数, $\therefore E(X|Y)$ 是一个随机变量. 用类似的方法可以解释书上第 7 章的式 (6.2).

例 11.1 (书上第 7 章的式 (6.2))**.** 对于 $x\in\mathbb{N}_+$, 有

$$x=\sum_{k=1}^{\infty} I[x\ge k].$$

将 x 替换为 $\Omega\to\mathbb{N}_+$ 的随机变量 X, 则有

$$X=\sum_{k=1}^{\infty} I[X\ge k].$$

12 第 6 章习题

6.1 (a)**.** 设 $N_1,N_2,\cdots$ 独立同分布, $N_1\sim\mathcal{B}(8,0.6)$, $M\sim\mathcal{P}(10)$, M 与 $\{N_j\}$ 独立. 对于 $T=\sum\limits_{j=1}^{M}N_j, k\ge 1$, 给出 $T|\{M=k\}$ 的概率分布.

解. N_j 表示成功概率为 0.6 的 8 次独立重复试验中成功的次数. $\because N_1,N_2,\cdots$ 独立, $\therefore T|\{M=k\}=N_1+\cdots+N_k$ 表示成功概率为 0.6 的 $8k$ 次独立重复试验中成功的次数, 服从二项分布 $\mathcal{B}(8k,0.6)$.

6.2. 设 $N\sim\mathcal{B}(8,0.6)$, $M\sim\mathcal{P}(10)$, M 与 N 独立. 如果一棋手将参加 M 场比赛, 每场比赛预计下 N 盘棋, 他期望总共能下多少盘棋?

解. 设这个棋手一共要下 X 盘棋. 与 6.1 类似, 有 $X|\{M=k\}\sim\mathcal{B}(8k,0.6)$. $\therefore E(X|M=k)=4.8k$. 由书上的定理 1.1 (3) 得

$$\begin{aligned}EX&=\sum_{k=0}^{\infty}E(X|M=k)P(M=k)\\&=\sum_{k=0}^{\infty}4.8k\cdot\frac{\lambda^k}{k!}e^{-\lambda}\\&=4.8\lambda e^{-\lambda}\sum_{k=1}^{\infty}\frac{\lambda^{k-1}}{(k-1)!}\\&=4.8\lambda e^{-\lambda}e^{\lambda}=4.8\lambda.\end{aligned}$$

把 $\lambda=10$ 代入得 $EX=48$.

6.3. 设 N 服从参数为 p 的几何分布, $X_1, X_2, \cdots$ 相互独立, 都服从参数为 λ 的指数分布, 当 N 和 $\{X_j\}$ 独立时, 计算 $W = \sum\limits_{j=1}^{N} X_j$ 的数学期望.

解. 由引理 10.1, $W|\{N=n\}$ 有概率密度

$$f_W(w;n) = \frac{\lambda^n}{(n-1)!} e^{-\lambda w} w^{n-1}, \quad w > 0.$$

$\therefore$

$$E(W|N=n) = \int_0^\infty w f_W(w;n) \mathrm{d}w = \int_0^\infty \frac{\lambda^n}{(n-1)!} e^{-\lambda w} w^n \mathrm{d}w.$$

令 $t = \lambda w$, 得

$$\begin{aligned} \int_0^\infty \frac{\lambda^n}{(n-1)!} e^{-\lambda w} w^n \mathrm{d}w &= \frac{1}{\lambda(n-1)!} \int_0^\infty e^{-t} t^n \mathrm{d}t \\ &= \frac{1}{\lambda(n-1)!} \Gamma(n+1) = \frac{n}{\lambda}. \end{aligned}$$

由书上的定理 1.1 (3) 得

$$\begin{aligned} EW &= \sum_{n=1}^{\infty} E(W|N=n) P(N=n) \\ &= \sum_{n=1}^{\infty} \frac{n}{\lambda} q^{n-1} p \\ &= \frac{p}{\lambda} \left(\frac{\mathrm{d}}{\mathrm{d}x} \sum_{n=0}^{\infty} x^n \right) \Bigg|_{x=q} \\ &= \frac{p}{\lambda} \left(\frac{1}{(1-x)^2} \right) \Bigg|_{x=q} = \frac{1}{\lambda p}. \end{aligned}$$

6.4. 某出租车在一天内遇到的红灯数 N 满足参数为 λ 的 Poisson 分布. 如果在每个红灯处的等待时间相互独立, 都在 $(0,2)$ 中均匀分布, 计算他一天内用于等候红灯的时间 X 的数学期望和方差.

解. 设 $X_1, X_2, \cdots$ 是在第 i 个红灯处的等待时间, 则 $X|\{N=n\} = X_1 + X_2 + \cdots + X_n$, X_i 独立同分布, $X_1 \sim \mathcal{U}(0,2)$, N 与 $\{X_i\}$ 独立. 有

$$E(X|N=n) = EX_1 + EX_2 + \cdots + EX_n = nEX_1 = n,$$

$$
\begin{aligned}
E(X^2|N=n) &= E(X_1+X_2+\cdots+X_n)^2 \\
&= E\left(X_1^2+X_2^2+\cdots+X_n^2+\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} X_iX_j\right) \\
&= nEX_1^2+\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} EX_iEX_j \\
&= \frac{4n}{3}+n(n-1).
\end{aligned}
$$

由书上的定理 1.1 (1) 得

$$
EX=\sum_{n=0}^{\infty}P(N=n)E(X|N=n)=\lambda,
$$

$$
\begin{aligned}
EX^2 &= \sum_{n=0}^{\infty}P(N=n)E(X^2|N=n) \\
&= \sum_{n=0}^{\infty}\left(\frac{4n}{3}+n(n-1)\right)\frac{\lambda^n}{n!}e^{-\lambda} \\
&= \frac{4}{3}\sum_{n=1}^{\infty}\frac{\lambda^n}{(n-1)!}e^{-\lambda}+\sum_{n=2}^{\infty}\frac{\lambda^n}{(n-2)!}e^{-\lambda} \\
&= \frac{4}{3}\lambda+\lambda^2.
\end{aligned}
$$

$\therefore$

$$
\mathrm{Var}(X)=EX^2-(EX)^2=\frac{4}{3}\lambda.
$$

6.9. 设 (X,Y) 有联合密度 $f(x,y)=ye^{-y(x+1)}, x>0, y>0$. 计算 $X|\{Y=y\}, Y|\{X=x\}$ 的概率密度 $f_{X|Y}(x|y), f_{Y|X}(y|x)$.

解. $f_{X|Y}(x|y), f_{Y|X}(y|x)$ 可以由 $f(x,y)$ 除以一个归一化因子得到. 有

$$
f_{X|Y}(x|y)=\frac{f(x,y)}{\int_0^\infty f(x,y)\mathrm{d}x}=\frac{ye^{-y(x+1)}}{e^{-y}}=ye^{-yx},
$$

$$
f_{Y|X}(y|x)=\frac{f(x,y)}{\int_0^\infty f(x,y)\mathrm{d}y}=(x+1)^2ye^{-y(x+1)}.
$$

6.14. 假设某设备的使用寿命 Y 服从 Weibull 分布, 有概率密度

$$
f(y)=aby^{b-1}e^{-ay^b}, \quad y\geq 0.
$$

对 $b=1/2$, 计算平均寿命 $E(Y-t|Y>t)$.

解. 由书上的定理 3.2,

$$\begin{aligned}E(Y|Y>t) &= \frac{E(YI[Y>t])}{P(Y>t)} \\ &= \frac{\int_t^\infty yf(y)\mathrm{d}y}{\int_t^\infty f(y)\mathrm{d}y} \\ &= \frac{ab\int_t^\infty y^{1/2}e^{-ay^{1/2}}\mathrm{d}y}{ab\int_t^\infty y^{-1/2}e^{-ay^{1/2}}\mathrm{d}y} \\ &= \frac{\int_t^\infty y^{1/2}e^{-ay^{1/2}}\mathrm{d}y}{\int_t^\infty y^{-1/2}e^{-ay^{1/2}}\mathrm{d}y}.\end{aligned}$$

令 $u = ay^{1/2}$, 则 $y = (u/a)^2, \mathrm{d}y = \dfrac{2}{a^2}u\mathrm{d}u$. 有

$$\begin{aligned}\frac{\int_t^\infty y^{1/2}e^{-ay^{1/2}}\mathrm{d}y}{\int_t^\infty y^{-1/2}e^{-ay^{1/2}}\mathrm{d}y} &= \frac{\int_{a\sqrt{t}}^\infty (u/a)e^{-u}(2/a^2)u\mathrm{d}u}{\int_{a\sqrt{t}}^\infty (u/a)^{-1}e^{-u}(2/a^2)u\mathrm{d}u} \\ &= \frac{\int_{a\sqrt{t}}^\infty u^2e^{-u}\mathrm{d}u}{a^2\int_{a\sqrt{t}}^\infty e^{-u}\mathrm{d}u} \\ &= \frac{e^{-a\sqrt{t}}(2+2a\sqrt{t}+a^2t)}{a^2e^{-a\sqrt{t}}} \\ &= \frac{2+2a\sqrt{t}+a^2t}{a^2}.\end{aligned}$$

$\therefore$

$$E(Y-t|Y>t) = E(Y|Y>t) - E(t|Y>t) = \frac{2+2a\sqrt{t}+a^2t}{a^2} - t.$$

6.15. 在某地任选一名出租车司机, 假设其开车经验 $Y \sim \Gamma(\alpha, \beta)$, 有概率密度

$$f_Y(y) = \frac{\beta^\alpha}{\Gamma(\alpha)}y^{\alpha-1}e^{-\beta y}, \quad y > 0.$$

如果开车经验为 y 的司机在一周内的收入 (单位: 元) 为 $N \sim \mathcal{P}(y)$, 已知一个司机在一周内的收入为 n 元, 计算他的开车经验的概率密度.

解.

6.17. 一台计算机有 7 个 USB 接口, 至少有 5 个接口正常时, 就认为计算机的 USB 接口能用. 假设开始时每个 USB 接口都正常, 其寿命 $X_1, X_2, \cdots, X_7$ 相互独立, 都服从参数为 λ 的指数分布. 计算这台计算机的 USB 能用的时间 Y 的概率分布.

解. 如果 $Y = y$, 则有两个 USB 接口在 y 时刻以前损坏, 有一个 USB 接口在 y 时刻损坏, 其余的 USB

接口在 y 时刻或以后损坏. $\therefore$

$$\begin{aligned}P(Y=y)&=\frac{7!}{2!1!4!}P(X_1<y,X_2<y,X_3=y,X_4\geq y,X_5\geq y,X_6\geq y,X_7\geq y)\\&=105(P(X_1<y))^2P(X_1=y)(P(X_1\geq y))^4\\&=105\lambda e^{-2\lambda y}(1-e^{-\lambda y})^2\mathrm{d}y,\end{aligned}$$

$\therefore Y$ 的概率密度为 $105\lambda e^{-2\lambda y}(1-e^{-\lambda y})^2$.

6.20. 设 $X_1,X_2,\cdots,X_n$ 独立同分布, 都服从参数为 λ 的指数分布, 证明随机变量

$$Y_1=X_{(1)},\quad Y_k=X_{(k)}-X_{(k-1)},\quad k=2,3,\cdots,n$$

相互独立, 且 $Y_i\sim\mathcal{E}((n+1-i)\lambda)$.

证明. $\because Y_1,Y_2,\cdots,Y_n$ 是连续型随机变量的线性组合, $\therefore Y_1,Y_2,\cdots,Y_n$ 是连续型随机变量. 有

$$\begin{aligned}P(Y_1=y_1,Y_2=y_2,\cdots,Y_n=y_n)&=P(X_{(1)}=y_1,X_{(2)}=y_1+y_2,\cdots,X_{(n)}=y_1+y_2+\cdots+y_n)\\&=n!P(X_1=y_1)P(X_1=y_1+y_2)\cdots P(X_1=y_1+y_2+\cdots+y_n).\end{aligned}$$

令 $u_1=y_1,u_2=y_1+y_2,\cdots,u_n=y_1+y_2+\cdots+y_n$, 则

$$\frac{\partial(u_1,u_2,\cdots,u_n)}{\partial(y_1,y_2,\cdots,y_n)}=\begin{vmatrix}1&0&\cdots&0\\1&1&\cdots&0\\\vdots&\vdots&\ddots&\vdots\\1&1&\cdots&1\end{vmatrix}=1,\quad\frac{\partial(u_1,u_2,\cdots,u_n)}{\partial(y_1,y_2,\cdots,y_n)}=1.$$

$$\begin{aligned}&n!P(X_1=y_1)P(X_1=y_1+y_2)\cdots P(X_1=y_1+y_2+\cdots+y_n)\\&=n!\lambda e^{-\lambda y_1}\cdot\lambda e^{-\lambda y_1+y_2}\cdots\lambda e^{-\lambda(y_1+y_2+\cdots+y_n)}\mathrm{d}u_1\mathrm{d}u_2\cdots\mathrm{d}u_n\\&=n\lambda e^{-n\lambda y_1}\cdot(n-1)\lambda e^{-(n-1)\lambda y_2}\cdots2\lambda e^{-2\lambda y_{n-1}}\cdot\lambda e^{-\lambda y_n}\mathrm{d}y_1\mathrm{d}y_2\cdots\mathrm{d}y_n.\end{aligned}$$

设 Y_i 的概率密度为 $f_{Y_i}(y_i)$, $Y_1,Y_2,\cdots,Y_n$ 的联合密度为 $f(y_1,y_2,\cdots,y_n)$, 则

$$\begin{aligned}f_{Y_i}(y_i)&=\int\cdots\int_{\mathbb{R}_+^{n-1}}P(Y_1=y_1,Y_2=y_2,\cdots,Y_n=y_n)\mathrm{d}y_1\cdots\mathrm{d}y_{i-1}\mathrm{d}y_{i+1}\cdots\mathrm{d}y_n\\&=(n-i+1)\lambda e^{-(n-i+1)\lambda y_i}\int_0^\infty n\lambda e^{-n\lambda y_1}\mathrm{d}y_1\cdots\int_0^\infty(n-i+2)\lambda e^{-(n-i+2)\lambda y_{i-1}}\mathrm{d}y_{i-1}\\&\quad\cdot\int_0^\infty(n-i)\lambda e^{-(n-i)\lambda y_{i+1}}\mathrm{d}y_{i+1}\cdots\int_0^\infty\lambda e^{-\lambda y_n}\mathrm{d}y_n\\&=(n-i+1)\lambda e^{-(n-i+1)\lambda y_i}.\end{aligned}$$

$\therefore Y_i\sim\mathcal{E}((n+1-i)\lambda)$.

$\because$

$$P(Y_1=y_1,Y_2=y_2,\cdots,Y_n=y_n)=f_{Y_1}(y_1)f_{Y_2}(y_2)\cdots f_{Y_n}(y_n)\mathrm{d}y_1\mathrm{d}y_2\cdots\mathrm{d}y_n,$$

$\therefore$

$$f(y_1, y_2, \cdots, y_n) = f_{Y_1}(y_1) f_{Y_2}(y_2) \cdots f_{Y_n}(y_n).$$

$\therefore Y_1, Y_2, \cdots, Y_n$ 相互独立. $\square$

6.21. 设取非负整数值的随机变量 X 有母函数 $g(s)$, 对非负整数 a, b, 求 $Y = aX + b$ 的母函数 $h(s)$.

解. Y 的取值为 $\{ja + b | j \in \mathbb{R}\}$. 有

$$\begin{aligned} g(s) &= \sum_{j=0}^{\infty} s^j P(X = j) \\ &= \sum_{j=0}^{\infty} (s^{1/a})^{ja} P(Y = ja + b) \\ &= (s^{1/a})^{-b} \sum_{j=0}^{\infty} (s^{1/a})^{ja+b} P(Y = ja + b) \\ &= s^{-b/a} h(s^{1/a}). \end{aligned}$$

$\therefore$

$$h(s^{1/a}) = g(s) s^{b/a} \Rightarrow h(s) = g(s^a) s^b.$$

6.24. 两人各抛均匀的硬币 n 次, 计算甲的正面次数恰好大于乙的正面次数 k 次的概率.

解. 设甲的正面次数为 X_1, 乙的正面次数为 X_2, $K = X_1 - X_2$, 则事件 "甲的正面次数恰好大于乙的正面次数 k 次" 为 $\{K = k\}$.

$X_1, X_2 \sim \mathcal{B}(n, 0.5)$, 有母函数

$$g_{X_1}(s) = g_{X_2}(s) = \frac{1}{2^n}(s+1)^n.$$

有

$$\begin{aligned}\sum_{i=-n}^{n} s^n P(K=i) &= E(s^K) \\ &= E(s^{X_1-X_2}) \\ &= E(s^{X_1}s^{-X_2}) \\ &= E(s^{X_1})E(s^{-X_2}) \\ &= E(s^{X_1})E((s^{-1})^{X_2}) \\ &= g_{X_1}(s)g_{X_2}(s^{-1}) \\ &= \frac{1}{2^{2n}}(s+1)^n\left(\frac{1}{s}+1\right)^n \\ &= \frac{1}{2^{2n}s^n}(s+1)^{2n} \\ &= \frac{1}{2^{2n}}\sum_{i=-n}^{n} C_{2n}^{i} s^{-n+i}.\end{aligned}$$

考察上式两边的 s^k 项, 有

$$P(K=k) = \frac{C_{2n}^{k+n}}{2^{2n}}.$$

注意. K 的取值可能是负数, $\therefore K$ 没有母函数.

第 6.29, 6.30 题需要用到下面的引理, 书上第 7 章的定理 3.1 也可以用这个引理来证明一般的情形.

引理 12.1. 若随机变量 X 的特征函数为 $\phi_X(t)$, 则随机变量 $Y=a(X+b)$ 的特征函数为 $\phi_Y(t)=\phi_X(at)e^{iabt}$.

证明. 由 $Y=a(X+b)$ 得 $X=a^{-1}Y-b$. 有

$$\phi_X(t)=Ee^{itX}=Ee^{it(a^{-1}Y-b)}=E(e^{it(a^{-1}Y)}e^{-itb})=e^{-itb}Ee^{i(ta^{-1})Y},\quad t\in\mathbb{R}.$$

令 $s=ta^{-1}$, 则 $t=as$,

$$\phi_X(as)=e^{-iabs}Ee^{isY}\Rightarrow\phi_Y(s)=Ee^{isY}=\phi_X(as)e^{iabs},\quad s\in\mathbb{R}.$$ □

6.29. 设 $X_1,X_2,\cdots$ 独立同分布, 都在 $(-1,1)$ 中均匀分布. 定义

$$Y_n=\sqrt{\frac{3}{n}}\sum_{j=1}^{n}X_j.$$

(1) 计算 X_1 的特征函数 $\phi_X(t)$;

(2) 计算 Y_n 的特征函数 $\phi_{Y_n}(t)$, 证明 $Y_n\xrightarrow{d}N(0,1)$.

证明. (1) $\because X_1$ 在 $(-1,1)$ 中均匀分布, $\therefore$

$$E\cos(tX)=\int_{-1}^{1}\frac{1}{2}\cos(tx)\mathrm{d}x=\frac{\sin t}{t},\quad E\sin(tX)=\int_{-1}^{1}\frac{1}{2}\sin(tx)\mathrm{d}x=0.$$

$\therefore$

$$\phi_X(t)=E\cos(tX)+iE\sin(tX)=\frac{\sin t}{t}.$$

(2) 有

$$\frac{Y_n}{\sqrt{3/n}}=\sum_{j=1}^{n}X_j.$$

令 $S_n=\sum\limits_{j=1}^{n}X_j$. 由书上的定理 6.2 (2) 得 S_n 的特征函数为

$$\phi_{S_n}(t)=\frac{\sin^n t}{t^n}.$$

由引理 12.1 得

$$\phi_{Y_n}(t)=\phi_{S_n}(t\sqrt{3/n})=\left(\frac{\sin(t\sqrt{3/n})}{t\sqrt{3/n}}\right)^n.$$

$\dfrac{\sin(t\sqrt{3/n})}{t\sqrt{3/n}}$ 有 Taylor 展开式

$$\frac{\sin(t\sqrt{3/n})}{t\sqrt{3/n}}=1-\frac{3t^2/n}{3!}+o((t/\sqrt{3/n})^3)=1-\frac{t^2}{2n}+o(1/n),$$

由书上的式 (6.15) 得

$$\phi_{Y_n}(t)=\left(1-\frac{t^2}{2n}+o(1/n)\right)^n\to e^{-t^2/2},\quad t\in\mathbb{R}.$$

由书上的定理 6.3 得 $Y_n\xrightarrow{d}N(0,1)$. □

6.30. 设 $X_1,X_2,\cdots$ 独立同分布, $X_1\sim\mathcal{E}(\lambda)$. 定义

$$S_n=\sum_{j=1}^{n}X_j,\quad Y_n=\frac{S_n-n/\lambda}{\sqrt{n/\lambda^2}}.$$

(1) 计算 X_1 的特征函数 $\phi_X(t)$;
(2) 计算 Y_n 的特征函数 $\phi_{Y_n}(t)$;
(3) 证明 $Y_n\xrightarrow{d}N(0,1)$.

证明. (1) 有

$$\phi_X(t)=Ee^{itX}=\int_0^\infty e^{itx}\cdot\lambda e^{-\lambda x}\mathrm{d}x=\frac{\lambda}{\lambda-it}.$$

(2) 由书上的定理 6.2 (2) 得 S_n 的特征函数为

$$\phi_{S_n}(t)=\left(\frac{\lambda}{\lambda-it}\right)^n.$$

由引理 12.1 得

$$\begin{aligned}\phi_{Y_n}(t) &= \phi_{S_n}(t/\sqrt{n/\lambda^2})e^{-it(n/\lambda)/\sqrt{n/\lambda^2}}\\ &= \left(\frac{1}{1-it/\sqrt{n}}\right)^n (e^{-it/\sqrt{n}})^n\\ &= \left(\frac{e^{-it/\sqrt{n}}}{1-it/\sqrt{n}}\right)^n.\end{aligned}$$

$\because$

$$\begin{aligned}e^z &= 1+z+\frac{z^2}{2}+o(|z|^2)\\ &= 1+z+\frac{z^2+z^3}{2}+o(|z|^2)\\ &= (1+z)\left(1+\frac{z^2}{2}+o(|z|^2)\right), \quad \forall z\in\mathbb{C},\end{aligned}$$

$\therefore$

$$\frac{e^{-it/\sqrt{n}}}{1-it/\sqrt{n}} = 1+\frac{(-it/\sqrt{n})^2}{2}+o(|-it/\sqrt{n}|^2) = 1-\frac{t^2}{2n}+o(1/n).$$

由书上的定理 6.3 得 $Y_n \xrightarrow{d} N(0,1)$. □

6.34. 对 $n=1,2,\cdots$, 设 $X_n\sim\mathcal{B}(n,p_n)$, $\lim\limits_{n\to\infty} np_n=\lambda$. 证明 $X_n\xrightarrow{d}\mathcal{P}(\lambda)$.

证明. 由书上的式 (6.5) 得 X_n 的分布函数为

$$\phi_{X_n}(t) = (1-p_n(1-e^{it}))^n = \left(1+\frac{\lambda(e^{it}-1)}{n}\right)^n.$$

$\because$

$$\lim_{n\to\infty}\phi_{X_n}(t) = \exp(\lambda(e^{it}-1)),$$

由书上的定理 6.3 得 $X_n\xrightarrow{d}\mathcal{P}(\lambda)$. □

6.35. 设 $X_1,X_2,\cdots,X_n$ 独立同分布, $S_k=X_1+X_2+\cdots+X_k, k\le n$. 如果 $E(|X_1|\big|S_n)$ 存在, 计算 $E(S_k|S_n)$.

解. $\because E(|X_1|\big|S_n)$ 存在, $\therefore E(X_1\big|S_n)$ 存在. 由书上的例 3.8 得 $E(X_1\big|S_n)=S_n/n$. $\therefore$

$$\begin{aligned}E(S_k|S_n) &= E(X_1+X_2+\cdots+X_k|S_n)\\ &= E(X_1\big|S_n)+E(X_2\big|S_n)+\cdots+E(X_n\big|S_n)\\ &= kE(X_1\big|S_n) = \frac{kS_n}{n}.\end{aligned}$$

13 第 7 章笔记

Markov 不等式中的不等号可以全部换成严格的不等号, 要证明这个结论, 只需将原证明的不等号全部换成严格的不等号即可.

设 $\xi_1, \xi_2, \cdots$ 是随机变量, 如果 $\xi_n \xrightarrow{p} 0$, 则记为 $\xi_n = o_p(1)$.

书上的定理 2.2 说明依概率收敛比几乎处处收敛要弱. 与弱大数律类似, 把强相合估计中的几乎处处收敛换成依概率收敛, 可以得到 (弱) 相合估计的定义.

定义 13.1. 设 w 是参数, w_n 是 w 的估计量, 如果当 $n \to \infty$ 时, $w_n \xrightarrow{p} w$, 则称 w_n 是 w 的**相合估计**.

相合估计不一定比不相合估计要好, 比如给线性回归的损失函数加上 L1 正则化会使得对参数的估计是不相合的, 但是加上 L1 正则化有很多好处.

Lindeberg-Feller 定理的直观理解为:

- 书上的式 (5.2) 的分母中有 $B_n^2 \to \infty (n \to \infty)$, 满足 Lindeberg 条件意味着 X_j 的随机性不能减小得太快, 否则 $\sum_{j=1}^{n} \frac{1}{B_n}(X_j - \mu_j)$ 会收敛到常数;
- 书上的式 (5.2) 的分子中有 $E((X_j - \mu_j)^2 I[|X_j - \mu_j| \geq \varepsilon B_n])$, 这当 $|X_j - \mu_j|$ 足够大时为 X_j 的方差. 如果对 $\forall j, |X_j - \mu_j| \geq \varepsilon B_n$, 那么 $E((X_j - \mu_j)^2 I[|X_j - \mu_j| \geq \varepsilon B_n]) = B_n^2$, 式 (5.2) 不收敛. $\therefore$ 满足 Lindeberg 条件意味着不能有很多的 X_j 满足 $|X_j - \mu_j| \geq \varepsilon B_n$.

书上的推论 5.2 体现了这两个直观理解.

从概率空间的角度来定义随机变量的 a.s. 收敛性.

定义 13.2. 设 $(\Omega, \mathcal{F}, P)$ 是概率空间, 随机变量 $X, X_1, X_2, \cdots$ 是 Ω 上的函数, 则 $X_n \to X$ a.s. 可以定义为对 $\exists A \subset \Omega$ 满足 $P(A) = 1$, 函数列 $\{X_n(\omega)\}$ 在 A 上收敛到 $X(\omega)$.

定理 13.1. 随机变量 $X, X_1, X_2, \cdots$ 满足定义 13.2 定义的 $X_n \to X$ a.s. 当且仅当 $X, X_1, X_2, \cdots$ 满足书上第 2 节定义的 $X_n \to X$ a.s..

证明. $(\Rightarrow)$ $\because$ 函数列 $\{X_n(\omega)\}$ 在 A 上收敛到 $X(\omega)$, $\therefore A \subset \{X_n \to X\}$. $\therefore P(X_n \to X) \geq P(A) = 1$.

$(\Leftarrow)$ 取 $A = \{X_n \to X\}$ 得. □

书上的例 6.3 省略了一些东西.

例 13.1 (书上的例 6.3 的前半部分)**.** 设 $X \in \mathcal{U}(0,1)$. 定义

$$\xi_n = nI[X < 1/n], \quad n = 1, 2, \cdots.$$

$\because \forall x > 0, \exists N$, 当 $n > N$ 时有 $x > 1/n$, $\therefore$ 如果 $X = x > 0$, 那么 $\exists N$, 当 $n > N$ 时有 $\xi_n = 0$. $\therefore$ 函数列 $\xi_n(\omega)$ 在 $\{X > 0\}$ 上收敛到 0. $\because P(X > 0) = 1$, 由定义 13.2 得 $\xi_n \to 0$ a.s.

补充书上定理 6.3 证明的一些细节. 下面所有的收敛都是 a.s. 收敛.

定理 13.2. 沿用书上定理 6.3 的符号. 有: $\tilde{X}_n$ 单调不减收敛到 $\tilde{Y}$.

证明. 当 $X_n \le M$ 时 $\tilde{X}_n = X_n$ 单调不减, 当 $X_n > M$ 时 $\tilde{X}_n = M$, $\therefore \tilde{X}_n$ 单调不减.

如果 $Y \le M$, 那么 $X_n \le Y \le M$, 这时有 $\tilde{X}_n = X_n, \tilde{Y} = Y$, $\therefore \tilde{X}_n$ 收敛到 $\tilde{Y}$.

如果 $Y > M$, $\because X_n$ 单调不减收敛到 Y, $\therefore \exists N$, 当 $n > N$ 时 $X_n > M$. 这时 $\tilde{X}_n = M = Y$. $\therefore \tilde{X}_n$ 收敛到 $\tilde{Y}$. □

有界收敛定理 (书上的定理 6.4) 和控制收敛定理 (书上的定理 6.6) 给出了从依概率收敛推出 L^1 收敛的一个条件: 随机变量序列有界或被一个期望有界的随机变量控制.

14 第 7 章习题

7.2. 证明: 设非负随机变量 X 有概率密度 $f(x)$, 则 $\forall$ 正数 M, 有

$$P(X \ge M) \le \frac{1}{M}\int_M^\infty x f(x)\mathrm{d}x.$$

证明. $\because X$ 有概率密度 $f(x)$, $\therefore$

$$\begin{aligned} MP(X \ge M) &= M\int_M^\infty f(x)\mathrm{d}x \\ &= \int_M^\infty Mf(x)\mathrm{d}x \\ &\le \int_M^\infty xf(x)\mathrm{d}x. \end{aligned}$$

□

7.3. 证明: 设非负随机变量 X 有分布函数 $F(x)$, 则 $\forall$ 正数 M, 有

$$P(X > M) \le \frac{\sqrt{(1-F(M))EX^2}}{M}.$$

证明. $\because X$ 的取值非负, $\therefore \forall M > 0, P(X > M) = P(X^2 > M^2)$. 由 Markov 不等式,

$$P(X > M) = P(X^2 > M^2) < \frac{EX^2}{M^2}.$$

$\therefore$

$$(P(X > M))^2 \le \frac{P(X > M)EX^2}{M^2} = \frac{(1-F(M))EX^2}{M^2}, \tag{14.1}$$

这里取等号是因为 $P(X > M)$ 可能 $= 0$. $\therefore$

$$P(X > M) \le \frac{\sqrt{(1-F(M))EX^2}}{M}.$$

□

7.4. 设非负随机变量 $X_1, X_2, \cdots, X_n$ 独立同分布, 对 $X_{(1)} = \min\{X_1, X_2, \cdots, X_n\}$, 证明

$$P(X_{(1)} \geq M) \leq \left(\frac{EX_1^2 P(X_1 \geq M)}{M^2}\right)^{n/2}.$$

证明. 与式 (14.1) 类似, 由 Markov 不等式得

$$P(X_1 \geq M) \leq \left(\frac{P(X_1 \geq M)EX_1^2}{M^2}\right)^{1/2}.$$

$\because X_1, X_2, \cdots, X_n$ 独立同分布, $\therefore$

$$P(X_{(1)} \geq M) = (P(X_1 \geq M))^n \leq \left(\frac{P(X_1 \geq M)EX_1^2}{M^2}\right)^{n/2}.$$ □

7.5. 设 X 有概率密度 $f(x) = \frac{x^m}{m!}e^{-x}$, $x \geq 0$, 计算 $EX, \mathrm{Var}(X)$, 证明

$$P(0 < X < 2(m+1)) \geq \frac{m}{m+1}.$$

证明. 有

$$EX = \int_0^\infty xf(x)\mathrm{d}x = \frac{1}{m!}\int_0^\infty x^{m+1}e^{-x}\mathrm{d}x = \frac{\Gamma(m+2)}{m!} = m+1,$$
$$EX^2 = \int_0^\infty x^2 f(x)\mathrm{d}x = \frac{1}{m!}\int_0^\infty x^{m+2}e^{-x}\mathrm{d}x = \frac{\Gamma(m+3)}{m!} = (m+1)(m+2),$$
$$\mathrm{Var}(X) = EX^2 - (EX)^2 = m+1.$$

由 Chebyshev 不等式,

$$P(|X-(m+1)| \geq (m+1)) \leq \frac{m+1}{(m+1)^2} = \frac{1}{m+1},$$

$\therefore$

$$P(0 < X < 2(m+1)) = 1 - P(|X-(m+1)| \geq (m+1)) = \frac{m}{m+1}.$$ □

7.6. 用书上第 2.4 节的例 4.3 的方法调查回答 "是" 的概率 p_1 和服用兴奋剂的概率 p, 由书上第 2.4 节的例 4.3 得 $p = 2p_1 - 1$. 设一次试验调查 n 个人, 随机变量 $\hat{p}_1$ 是 n 个人中回答 "是" 的比例, $\hat{p} := 2\hat{p}_1 - 1$.

(a) 计算 $E\hat{p}_1, \mathrm{Var}(\hat{p}_1), E\hat{p}, \mathrm{Var}(\hat{p})$;

(b) 对 $\varepsilon > 0$, 证明

$$P(|\hat{p}_1 - p_1| \geq \varepsilon) \leq \frac{1-p^2}{4n\varepsilon^2}, \quad P(|\hat{p} - p| \geq \varepsilon) \leq \frac{1-p^2}{n\varepsilon^2}.$$

证明. (a) 由书上的例 2.1 得 $E\hat{p}_1 = p_1, E\hat{p} = p$.

引入随机变量

$$X_j = \begin{cases} 1, & \text{第}j\text{个人回答“是”}, \\ 0, & \text{第}j\text{个人回答“否”}, \end{cases} \quad j = 1, 2, \cdots,$$

则 $X_1, X_2, \cdots$ 独立同分布, 且

$$P(X_j = 1) = p_1, \quad EX_j = EX_j^2 = p_1, \quad \hat{p}_1 = \frac{1}{n}\sum_{j=1}^{n} X_j.$$

$\because X_1, X_2, \cdots$ 独立, $\therefore$

$$\begin{aligned} E\hat{p}_1^2 &= E\left(\frac{1}{n}\sum_{j=1}^{n} X_j\right)^2 \\ &= \frac{1}{n^2}E\left(\sum_{i=1}^{n} X_i^2 + \sum_{i=1}^{n}\sum_{\substack{j=1 \\ j\neq i}}^{n} X_i X_j\right) \\ &= \frac{1}{n^2}\left(\sum_{i=1}^{n} EX_i^2 + \sum_{i=1}^{n}\sum_{\substack{j=1 \\ j\neq i}}^{n} EX_i EX_j\right) \\ &= \frac{1}{n^2}(np_1 + n(n-1)p_1^2) \\ &= \frac{p_1 + (n-1)p_1^2}{n}. \end{aligned}$$

$\therefore$

$$\mathrm{Var}(\hat{p}_1) = E\hat{p}_1^2 - (E\hat{p}_1)^2 = \frac{p_1(1-p_1)}{n} = \frac{1-p^2}{4n},$$

$$\mathrm{Var}(\hat{p}) = \mathrm{Var}(2\hat{p}_1 + 1) = 4\,\mathrm{Var}(\hat{p}_1) = \frac{1-p^2}{n}.$$

(b) 由 Chebyshev 不等式得

$$P(|\hat{p}_1 - p_1| \geq \varepsilon) = P(|\hat{p}_1 - E\hat{p}_1| \geq \varepsilon) \leq \frac{1}{\varepsilon^2}\,\mathrm{Var}(\hat{p}_1) = \frac{1-p^2}{4n\varepsilon^2},$$

$$P(|\hat{p} - p| \geq \varepsilon) = P(|\hat{p} - E\hat{p}| \geq \varepsilon) \leq \frac{1}{\varepsilon^2}\,\mathrm{Var}(\hat{p}) = \frac{1-p^2}{n\varepsilon^2}.$$

□

7.7 (有修改)**.** 对随机变量 X 进行观测, 设第 j 次观测的结果为 X_j, $X_1, X_2, \cdots$ 独立同分布, 且与 X 同分布. 书上的例 2.2 给出了 $\mu = EX, \sigma^2 = \mathrm{Var}(X), F(x) = P(X \leq x)$ 的强相合估计

$$\hat{\mu}_n = \frac{1}{n}\sum_{j=1}^{n} X_j, \quad \hat{\sigma}_n^2 = \frac{1}{n-1}\sum_{j=1}^{n}(X_j - \hat{\mu}_n)^2, \quad F_n(x) = \frac{1}{n}\sum_{j=1}^{n} I[X_j \leq x],$$

(1) 求 $E\hat{\mu}_n, E\hat{\sigma}_n^2, EF_n(x)$.

(2) 给定 x, n, 求 $\operatorname{Var} F_n(x)$.

解. (1) 有

$$E\hat{\mu}_n = \frac{1}{n}\sum_{j=1}^{n} EX_j = \frac{1}{n}\sum_{j=1}^{n}\mu = \mu,$$

$$E(F_n(x)) = \frac{1}{n}\sum_{j=1}^{n} EI[X_j \le x] = \frac{1}{n}\sum_{j=1}^{n} P(X_j \le x) = \frac{1}{n}\sum_{j=1}^{n} P(X \le x) = P(X \le x).$$

$\because$

$$\begin{aligned}
\hat{\sigma}_n^2 &= \frac{1}{n-1}\sum_{j=1}^{n}(X_j - \hat{\mu}_n)^2 \\
&= \frac{1}{n-1}\sum_{j=1}^{n}(X_j^2 - \hat{\mu}_n^2 + 2X_j\hat{\mu}_n) \\
&= \frac{1}{n-1}\left(\sum_{j=1}^{n} X_j^2 - n\hat{\mu}_n^2 + 2\hat{\mu}_n\sum_{j=1}^{n} X_j\right) \\
&= \frac{1}{n-1}\left(\sum_{j=1}^{n} X_j^2 - n\hat{\mu}_n^2 + 2n\hat{\mu}_n^2\right) \\
&= \frac{1}{n-1}\sum_{j=1}^{n} X_j^2 - \frac{n}{n-1}\hat{\mu}_n^2,
\end{aligned}$$

$\therefore$

$$\begin{aligned}
E\hat{\sigma}_n^2 &= \frac{1}{n-1}\sum_{j=1}^{n} EX_j^2 - \frac{n}{n-1}E\hat{\mu}_n^2 \\
&= \frac{1}{n-1}\sum_{j=1}^{n} EX_j^2 - \frac{n}{n-1}E\left(\frac{1}{n}\sum_{j=1}^{n} X_j\right)^2 \\
&= \frac{n}{n(n-1)}\sum_{j=1}^{n} EX_j^2 - \frac{\sum\limits_{i=1}^{n} EX_i^2 + \sum\limits_{i=1}^{n}\sum\limits_{\substack{j=1\\ j\ne i}}^{n} EX_iEX_j}{n(n-1)} \\
&= \frac{n-1}{n(n-1)}\sum_{j=1}^{n} EX^2 - \frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\ne i}}^{n}(EX)^2 \\
&= EX^2 - (EX)^2 = \operatorname{Var}(X).
\end{aligned}$$

(2) $nF_n(x)$ 服从二项分布 $\mathcal{B}(n, P(X_j \le x)) = \mathcal{B}(n, F(x))$. $\therefore$

$$\operatorname{Var}(F_n(x)) = \frac{1}{n^2}\operatorname{Var}(nF_n(x)) = \frac{1}{n^2}\cdot nF(x)(1-F(x)) = \frac{1}{n}F(x)(1-F(x)).$$

注. (1) $\because E\hat{\mu}_n = EX, E\hat{\sigma}_n^2 = \operatorname{Var}(X), EF_n(x) = F(x)$, $\therefore E\hat{\mu}_n, E\hat{\sigma}_n^2, EF_n(x)$ 都是无偏估计.

(2) 给定 x, 当 n 充分大时, 由中心极限定理得 $\dfrac{nF_n(x) - nF(x)}{\sqrt{nF(x)(1-F(x))}}$ 近似服从标准 Gauss 分布. 当 X 服从 Bernoulli 分布时, 可以得到书上的定理 4.1.

第 7.9, 7.10 题需要用到下面的引理.

引理 14.1. 设 $*: \mathbb{R}^2 \to \mathbb{R}$ 是某种二元运算, 如果对任意的有极限的数列 $\{a_n\}, \{b_n\}$, $\lim\limits_{n\to\infty}(a_n * b_n) = \lim\limits_{n\to\infty} a_n * \lim\limits_{n\to\infty} b_n$, 那么对 $X_n \to x$ a.s., $Y_n \to y$ a.s., $\lim\limits_{n\to\infty}(X_n * Y_n) = x * y$ a.s.

证明. $\because X_n \to x$ a.s., $\therefore P\left(\lim\limits_{n\to\infty} X_n = x\right) = 1$. 对称地, $P\left(\lim\limits_{n\to\infty} Y_n = y\right) = 1$. $\therefore$ 由

$$P\left(\lim_{n\to\infty} X_n = x \cup \lim_{n\to\infty} Y_n = y\right) = P\left(\lim_{n\to\infty} X_n = x\right) + P\left(\lim_{n\to\infty} Y_n = y\right) - P\left(\lim_{n\to\infty} X_n = x, \lim_{n\to\infty} Y_n = y\right)$$

得

$$\begin{aligned}
&P\left(\lim_{n\to\infty} X_n = x, \lim_{n\to\infty} Y_n = y\right) \\
&= P\left(\lim_{n\to\infty} X_n = x\right) + P\left(\lim_{n\to\infty} Y_n = y\right) - P\left(\lim_{n\to\infty} X_n = x \cup \lim_{n\to\infty} Y_n = y\right) \\
&= 1.
\end{aligned}$$

$\because \left\{\lim\limits_{n\to\infty}(X_n * Y_n) = x * y\right\} \supset \left\{\lim\limits_{n\to\infty} X_n = x, \lim\limits_{n\to\infty} Y_n = y\right\}$, $\therefore$

$$P\left(\lim_{n\to\infty}(X_n * Y_n) = x * y\right) \geq P\left(\lim_{n\to\infty} X_n = x, \lim_{n\to\infty} Y_n = y\right) = 1,$$

$\therefore P\left(\lim\limits_{n\to\infty}(X_n * Y_n) = x * y\right) = 1$, 即 $\lim\limits_{n\to\infty}(X_n * Y_n) = x * y$ a.s. □

7.9 (有修改)**.** 设 $X_0, X_1, \cdots$ 是独立同分布的随机变量序列, $\mu = EX_0$, 对非零常数 a, b, 定义

$$Y_k = aX_k + bX_{k-1} + c, \quad k = 1, 2, \cdots$$

计算 $n^{-1}\sum\limits_{k=1}^{n} Y_k$ 的 a.s. 极限.

解. 由强大数律,

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} X_k = \lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} X_k = EX_0 = \mu, \text{ a.s.}$$

在引理 14.1 中以 "+" 代 "$*$" 得

$$\begin{aligned}
\frac{1}{n}\sum_{k=1}^{n} Y_k &= \frac{1}{n}\left(a\sum_{k=1}^{n} X_k + b\sum_{k=0}^{n-1} X_k + nc\right) \\
&= a \cdot \frac{1}{n}\sum_{k=1}^{n} X_k + b \cdot \frac{1}{n}\sum_{k=0}^{n-1} X_k + c \\
&\to a\mu + b\mu + c \text{ a.s.} \quad (n \to \infty).
\end{aligned}$$

注. (1) 原题没有定义 X_0, 但 Y_1 的定义式中有 X_0, 所以这里改了一下题目, 加上了 X_0 的定义.

(2) $\because Y_k$ 是否独立不能确定, $\therefore$ 不能直接对 Y_k 用强大数律.

7.10. 设 $X_1, X_2, \cdots$ 独立同分布, 都在 $(0, \pi/2)$ 中均匀分布,

$$Y_n = \frac{\sin X_1 + \sin X_2 + \cdots + \sin X_n}{\cos X_1 + \cos X_2 + \cdots + \cos X_n}, \quad n \geq 1,$$

当 $n \to \infty$ 时, 计算 Y_n 的 a.s. 极限.

解. 有

$$E(\cos X_1) = \int_0^{\pi/2} \cos x \frac{2}{\pi} \mathrm{d}x = \frac{2}{\pi}, \quad E(\sin X_1) = \int_0^{\pi/2} \sin x \frac{2}{\pi} \mathrm{d}x = \frac{2}{\pi}.$$

$\therefore$ 当 $n \to \infty$ 时,

$$\frac{\sin X_1 + \sin X_2 + \cdots + \sin X_n}{n} \to E(\sin X_1) = \frac{2}{\pi}, \text{ a.s.}$$

$$\frac{\cos X_1 + \cos X_2 + \cdots + \cos X_n}{n} \to E(\cos X_1) = \frac{2}{\pi}, \text{ a.s.}$$

在引理 14.1 中以 “/” 代 “$*$” 得

$$Y_n = \frac{\dfrac{\sin X_1 + \sin X_2 + \cdots + \sin X_n}{n}}{\dfrac{\cos X_1 + \cos X_2 + \cdots + \cos X_n}{n}} \to \frac{E(\sin X_1)}{E(\cos X_1)} = 1, \text{ a.s.}$$

7.14. 设选民中赞同某候选人的比例是未知数 p, 已知 $p \in (0.01, 0.99)$. 对 p 进行调查, 为了以 99% 的把握对 p 的预测的绝对误差不超过 1%, 需要调查多少选民?

解. 设调查了 n 人, n 个人中赞同某候选人的人数为 S_i, 则 $S_i \sim \mathcal{B}(n, p)$. 事件 “调查了 n 人时, 对 p 的预测的绝对误差不超过 1%” 可以表示为

$$\{-0.01 \leq S_n/n - p \leq 0.01\} = \left\{-\frac{0.01n}{\sqrt{npq}} \leq \frac{S_n - np}{\sqrt{npq}} \leq \frac{0.01n}{\sqrt{npq}}\right\},$$

其中 $q = 1 - p$.

由平均值不等式, $\sqrt{pq} \leq \dfrac{p+q}{2} = \dfrac{1}{2}$ (当 $p = q$ 时取等号). $\because p \in (0.01, 0.99)$, $\therefore$ 等号可以取到, 有

$$\begin{aligned}\left\{-\frac{0.01n}{\sqrt{npq}} \leq \frac{S_n - np}{\sqrt{npq}} \leq \frac{0.01n}{\sqrt{npq}}\right\} &\subset \left\{-\frac{0.01n}{\sqrt{n}/2} \leq \frac{S_n - np}{\sqrt{npq}} \leq \frac{0.01n}{\sqrt{n}/2}\right\} \\ &= \left\{-0.02\sqrt{n} \leq \frac{S_n - np}{\sqrt{npq}} \leq 0.02\sqrt{n}\right\}.\end{aligned}$$

先假设 $np \geq 5, nq \geq 5$. 由中心极限定理得 $\dfrac{S_n - np}{\sqrt{npq}} \sim N(0,1)$. $\therefore$

$$P(-0.01 \leq S_n/n - p \leq 0.01) \leq P\left(-0.02\sqrt{n} \leq \frac{S_n - np}{\sqrt{npq}} \leq 0.02\sqrt{n}\right) = 2\Phi(0.02\sqrt{n}) - 1.$$

$\because$ 当 $n = 16641$ 时 $2\Phi(0.02\sqrt{n}) - 1 = 0.99$, $\therefore$ 需要调查 16641 个人. 有

$$np = 166.41 \geq 5, \quad nq = 166.41 \geq 5.$$

注. $n = 16641$ 是对于 p 可以取到 $\frac{1}{2}$ 而言的. 如果已知 $0.01 < p < 0.4$ 或 $0.6 < p < 0.09$, 那么只需要调查 3835 个人. p 越接近 1/2, 要确定 p 所需的信息量就越大.

7.15. 设独立同分布的随机变量 $X_1, X_2, \cdots, X_n$ 和独立同分布的随机变量 $Y_1, Y_2, \cdots, Y_m$ 相互独立, $EX_1 = \mu_1, \mathrm{Var}(X_1) = \sigma_1^2, EY_1 = \mu_2, \mathrm{Var}(Y_1) = \sigma_2^2$. 对充分大的 n, m, 求 $\frac{1}{n}\sum_{j=1}^{n} X_j - \frac{1}{m}\sum_{k=1}^{m} Y_k$ 的近似分布.

解. 由中心极限定理得

$$\frac{\sum_{j=1}^{n} X_j - n\mu_1}{\sqrt{n\sigma_1^2}} \xrightarrow{d} N(0,1), \quad \frac{\sum_{j=1}^{m} Y_j - n\mu_2}{\sqrt{m\sigma_2^2}} \xrightarrow{d} N(0,1).$$

$\therefore$

$$\frac{1}{n}\sum_{j=1}^{n} X_j \sim N(\mu_1, \sigma_1^2/n), \quad \frac{1}{m}\sum_{j=1}^{m} Y_j \sim N(\mu_2, \sigma_2^2/m).$$

$\because X_1, X_2, \cdots, X_n, Y_1, Y_2, \cdots, Y_m$ 相互独立, 由书上第 4 章的定理 1.1 (3) 得 $\frac{1}{n}\sum_{j=1}^{n} X_j$ 和 $\frac{1}{m}\sum_{j=1}^{m} Y_j$ 相互独立, 由第 4.27 题得 $\frac{1}{n}\sum_{j=1}^{n} X_j - \frac{1}{m}\sum_{k=1}^{m} Y_k$ 服从 Gauss 分布.

$\because$

$$E\left(\frac{1}{n}\sum_{j=1}^{n} X_j - \frac{1}{m}\sum_{k=1}^{m} Y_k\right) = E\left(\frac{1}{n}\sum_{j=1}^{n} X_j\right) - E\left(\frac{1}{m}\sum_{k=1}^{m} Y_k\right) = \mu_1 - \mu_2,$$

$$\mathrm{Var}\left(\frac{1}{n}\sum_{j=1}^{n} X_j - \frac{1}{m}\sum_{k=1}^{m} Y_k\right) = \mathrm{Var}\left(\frac{1}{n}\sum_{j=1}^{n} X_j\right) + \mathrm{Var}\left(-\frac{1}{m}\sum_{k=1}^{m} Y_k\right) = \frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m},$$

$\therefore$

$$\frac{1}{n}\sum_{j=1}^{n} X_j - \frac{1}{m}\sum_{k=1}^{m} Y_k \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}\right).$$

7.16. 某保险公司每天平均受理 18 份投保, 每份投保平均缴款 $\mu = 3.2$ 万元, 标准差为 $\sigma = 2.8$ 万元. 假定每天受理的投保数 N 服从 Poisson 分布, 每份投保的缴款相互独立, 且与 N 独立. 计算

(1) 明天投保缴款 S 万元的数学期望和方差;

(2) $P(S \leq 50)$.

解. (1) $\because N$ 服从 Poisson 分布, $\therefore E(N) = \mathrm{Var}(n) = 18$.

设明天第 i 份投保的缴款为 X_i, 则 $X_1, X_2, \cdots$ 独立同分布. 有 $E(S|N = n) = n\mu$, $\quad E(S|N) = N\mu$, $\therefore ES = E(E(S|N)) = \mu EN$.

有

$$E(S^2|N = n) = E\left(\left(\sum_{j=1}^{n} X_i\right)^2 \middle| N = n\right).$$

$\because N$ 与 $\{X_i\}$ 独立, 由书上第 4 章的定理 1.1 (3) 得 N 与 $\left(\sum_{j=1}^{n} X_i\right)^2$ 独立. $\therefore$

$$\begin{aligned} E\left(\left(\sum_{j=1}^{n} X_i\right)^2 \middle| N = n\right) &= E\left(\sum_{j=1}^{n} X_i\right)^2 \\ &= \operatorname{Var}\left(\sum_{j=1}^{n} X_i\right) - \left(E\left(\sum_{j=1}^{n} X_i\right)\right)^2 \\ &= n\sigma^2 + (n\mu)^2. \end{aligned}$$

$\therefore$

$$E(S^2|N) = N\sigma^2 + N^2\mu^2,$$

$$\begin{aligned} ES^2 &= E(N\sigma^2 + N^2\mu^2) \\ &= \sigma^2 EN + \mu^2 EN^2 \\ &= \sigma^2 EN + \mu^2(\operatorname{Var}(N) + (EN)^2) \\ &= \sigma^2 EN + \mu^2(EN + (EN)^2), \end{aligned}$$

$$\operatorname{Var}(S) = ES^2 - (ES)^2 = \sigma^2 EN + \mu^2 EN.$$

代入得 $ES = 57.6, \operatorname{Var}(S) = 325.44$.

(2) 可以把一天分成 m (m 足够大) 个不相交的区间, 使得所有区间内的受理的投保数的均值相同. 设 N_i 是第 i 个区间内受理的投保数, S_i 是第 i 个区间内受理的投保的缴款总数, 则 $N_i \sim \mathcal{P}(18/m), S = S_1 + S_2 + \cdots + S_m$.

由中心极限定理得 S 近似服从 Gauss 分布. $\therefore$

$$P(S \le 50) = P\left(\frac{S - ES}{\sqrt{\operatorname{Var}(S)}} \le \frac{50 - ES}{\sqrt{\operatorname{Var}(S)}}\right) \approx \Phi\left(\frac{50 - ES}{\sqrt{\operatorname{Var}(S)}}\right).$$

代入得 $P(S \le 50) \approx 0.3372$.

注意. $\because N$ 的值不确定, N 的值有可能很小, $\therefore$ 不能由于每天平均受理的投保数很多就认为 S 一定是很多次投保的缴款数之和, 近似服从 Gauss 分布.

7.20. 证明公式

$$\lim_{n\to\infty} e^{-n} \sum_{k=0}^{n-1} \frac{n^k}{k!} = \frac{1}{2}, \quad \lim_{n\to\infty} e^{-n} \sum_{k=0}^{n} \frac{n^k}{k!} = \frac{1}{2}.$$

证明. □

7.21. 如果 $\sqrt{n}\xi_n \xrightarrow{d} N(0,1)$, 证明 $\xi_n \xrightarrow{p} 0$.

证明. 设 ξ_n 的分布函数为 $F_n(x)$, 则 $\sqrt{n}\xi_n$ 的分布函数为

$$G_n(x) = P(\sqrt{n}\xi_n \le x) = P(\xi_n \le x/\sqrt{n}) = F_n(x/\sqrt{n}).$$

$\because \sqrt{n}\xi_n \xrightarrow{d} N(0,1)$, $\therefore \forall x \in \mathbb{R}$,

$$\lim_{n\to\infty} G_n(x) = \Phi(x).$$

$\because \forall \varepsilon > 0$,

$$\begin{aligned} P(|\xi_n - 0| \ge \varepsilon) &= 1 - P(-\varepsilon < \xi_n < \varepsilon) \\ &= 1 - F_n(\varepsilon) + F_n(-\varepsilon) \\ &= 1 - G_n(\varepsilon\sqrt{n}) + G_n(-\varepsilon\sqrt{n}) \end{aligned}$$

□

7.22. 设 $X_1, X_2, \cdots$ 相互独立, $X_j \sim \mathcal{B}(j, p_j), S_n = X_1 + X_2 + \cdots + X_n$. 当 $p_j \in [0.1, 0.9]$, 证明

$$\frac{S_n - ES_n}{\sqrt{\mathrm{Var}(S_n)}} \xrightarrow{d} N(0,1).$$

证明. 有 $\mu_j = EX_j = jp_j, \sigma_j^2 = \mathrm{Var}(X_j) = jp_j(1-p_j), B_n^2 = \sum_{j=1}^n \sigma_j^2$. $\because p_j \in [0.1, 0.9]$, $\therefore$

$$B_n^2 = \sum_{j=1}^n jp_j(1-p_j) \ge \sum_{j=1}^n 0.1 \cdot 0.9j = 0.09\frac{n(n+1)}{2}.$$

有

$$\max_{1\le j\le n} |X_j - \mu_j| < \max_{1\le j\le n} j = n \quad \text{a.s.}$$

$\because \lim_{n\to\infty} n/B_n$ □

7.23. 如果 $\xi_n \xrightarrow{d} N(0,1), \eta_n \xrightarrow{p} 0$, 证明 $\xi_n + \eta_n \xrightarrow{d} N(0,1)$.

证明. $\because \xi_n \xrightarrow{d} N(0,1)$, $\therefore \forall x$, $\lim_{n\to\infty} P(\xi_n \le x) = \Phi(x)$, 即 $\forall \varepsilon, \exists N$, 当 $n \ge N$ 时有 $|P(\xi_n \le x) - \Phi(x)| < \frac{\varepsilon}{2}$.

$\because$ □