概率论与数理统计（第二版）

黄炜 张帼奋 张奕 张彩伢 主编

教材课后题详解答案

前言

之前做了微积分卢兴江版的作业详解，也是有很多人来问我能不能做其他数学基础课的详解答案。当然这项工作其实历史上也有人简单的做过一下，但是一份 Latex 编辑过的电子版答案还是没有的。借着浙大版概统这本书已经到了第二版，我也就把天赋带到概统好了，编辑下不止一章不止两章而是这本书全部章节详解答案，毕竟虽然书后有着每道题的题目答案，但是只有数据没有过程还是对很多人来说产生理解上的问题。

当然这里只列出了课后习题中 A 和 B 组的题目详解答案，毕竟思考题部分最后参考答案部分也比较详细。概率论与数理统计这门课本身难度没有微积分高，但是因为会涉及到很多计算过程，加上还有分布律表格等等，生产答案速度应该也不比微积分快到哪里去。

这里要感谢合作编者萝卜和yeeting，感谢他们提供的大量习题解答，毕竟我一个弄微积分为主的人确实概统题目这一块需要时间恢复记忆，有一些答案辅助也是如虎添翼的。

也欢迎大家在公众号“路老师的nonsense collection”后台回复我订正答案，不日会更新答案错误。也希望各位尊重原作者，尊重版权。

也是又做了一点微小的工作，很惭愧，谢谢大家的支持。

可爱的路老师

2025 年 11 月于杭州

目录

第 1 章　概率论的基本概念

A1. 为了解吸烟时人体健康产生的影响, 对一社区居民进行抽样调查, 分别用 0, 1, 2 表示不吸烟、少量吸烟及吸烟较多, 再用 a, b, c 表示身体健康、一般及患病. 例如, $(0, a)$ 就表示抽到的居民是不吸烟的健康者.

(1) 问试验的样本空间共有多少个样本点?

(2) 设 A ={抽到的居民身体健康}, 写出 A 所包含的样本点;

(3) 设 B ={抽到的居民不吸烟}, 写出 B 所包含的样本点.

解:

A2. 写出下列随机试验中的随机事件 A 和 B 所包含的样本点:

(1) 掷一颗骰子,

事件 $A=\{$点数不大于2$\}$, 事件 $B=\{$点数大于3$\}$;

(2) 将一颗骰子掷两次,

事件 $A=\{$两次点数之差的绝对值为2$\}$,

事件 $B=\{$第一次点数是第二次点数的3 倍$\}$;

(3) 在以 $(0,0)$, $(1,0)$, $(0,1)$ 为顶点的三角形内随机取一点, 记其坐标 (x, y),

事件 $A=\{$横坐标 x 不大于纵坐标 $y\}$

事件 B ={横坐标 x 小于 0.5 且纵坐标 y 大于 0.5}.

解:

A3. 写出下列随机试验的样本空间:

(1) 一盒中有编号为 1,2,3 的三个球, 依次随机地抽取 2 次, 每次取 1 个, 无放回, 观察两个球的编号;

(2) 一盒中有编号为 1,2,3 的三个球, 依次随机地抽取 2 次, 每次取 1 个, 有放回, 观察两个球的编号;

(3) 某超市每天的营业时间为 7: 00 - 22: 00, 观察某天进入该超市的人数;

(4) 在以 (0,0), (1,0), (0,1) 为顶点的三角形内 (含边界) 随机取一点, 记录其坐标.

解:

A4. 设 A, B, C 为 3 个随机事件, 请用事件的运算关系式表示:

(1) A, B, C 至少有 2 个发生;

(2) A, B, C 最多有 1 个发生;

(3) A, B, C 恰有 1 个不发生;

(4) A, B, C 至少有 1 个不发生.

解:

A5. 设 A, B 为两个随机事件, 且 A, B 中至少有一个发生的概率为 0.9.

(1) 若 B 发生的同时 A 不发生的概率为 0.4, 求 $P(A)$;

(2) 若 $P(B) = 0.6$, 求 A 发生的同时 B 不发生的概率.

解:

A6. 设 A, B 为两个随机事件, 已知 $P(A) = 0.5$, $P(B) = 0.4$, 在下列两种情况下分别求 $P(A \cup B)$ 和 $P(A \cap B)$:

(1) A 与 B 互不相容;

(2) 当 B 发生时必有 A 发生.

解:

A7. 设事件 A 与 B 互不相容, $P(A) = 0.3$, $P(B) = 0.5$, 求:

(1) A 与 B 至少有一个发生的概率;

(2) A 与 B 都不发生的概率;

(3) A 不发生的同时 B 发生的概率.

解:

A8. 设 A, B 为两个随机事件, 已知 $P(A) = 0.5$, $P(B) = 0.4$, $P(A \cup B) = 0.6$, 求:

(1) $P(A\overline{B})$;

(2) $P(\overline{AB})$;

(3) $P(\overline{A} \cup \overline{B})$.

解:

A9. 已知一宿舍有 6 位学生, 其中有 2 位为统计学专业. 从该宿舍随机选 2 位学生, 求:

(1) 至少有 1 位是统计学专业的概率;

(2) 最多有 1 位是统计学专业的概率.

解:

A10. 同时掷两颗均匀的骰子,

(1) 求两颗骰子点数不同的概率;

(2) 求某一颗骰子点数是另一颗骰子点数 3 倍的概率;

(3) 已知 2 颗骰子的点数不同, 求某一颗骰子点数是另一颗骰子点数 3 倍的概率.

解:

A11. 假设一个人出生的月份在一年的 12 个月是等概率的, 求一宿舍 6 位学生中至少有 2 人生日在同一个月的概率.

解:

A12. 一袋中有 10 个球, 其中 8 个是红球, 每次摸一球, 共摸 2 次, 在有放回抽样与无放回抽样两种抽样方式下分别求:

(1) “两次均为红球”的概率;

(2) “恰有 1 个红球”的概率;

(3) “第 2 次是红球”的概率.

解:

A13. 在一个 30 人的班级中有两个“王姓”学生, 现将全班学生随机排成一排, 求:

(1) 两个 “王姓”学生正好中间隔一个位置的概率;

(2) 两个 “王姓”学生正好排在最中间的概率.

解:

A14. 一个盒子中有 7 个球, 其中 2 个红球, 3 个黑球, 2 个白球, 每次摸一球 (无放回), 共摸 3 次,求:

(1) 摸到的球恰是 1 红、1 黑、1 白的概率;

(2) 摸到的球全是黑球的概率;

(3) 第 1 次摸到红球、第 2 次摸到黑球且第 3 次摸到白球的概率.

解:

(3) 设该事件为 C, 则 $P(C)=\frac{2}{7}\times\frac{3}{6}\times\frac{2}{5}=\frac{2}{35}$.

A15. 编号为 $1,2,\cdots,9$ 的 9 辆车, 随机地停入编号为 $1,2,\cdots,9$ 的 9 个车位中. 当车号与车位编号一样时称该车配对, 求:

(1) 1 号车配对的概率;

(2) 1 号车配对而 9 号车不配对的概率.

解:

A16. 依次将 5 个不同的球随机放入 3 个不同的盒子中, 盒子容量不限, 求 3 号盒子中恰好有两个球的概率.

解:

A17. 设 A, B 为两个随机事件, 已知 $P(A)=0.5$, $P(B)=0.4$, $P(A\cup B)=0.6$,求:

(1) $P(A|B)$;

(2) $P(B|A)$;

(3) $P(\overline{A}|\overline{B})$.

(4) $P(\overline{B}|\overline{A})$.

解:

A18. 设 A, B 为两个随机事件, 已知 $P(A)=0.6$, $P(B|A)=0.4$, $P(A|B)=0.5$, 求:

(1) $P(AB)$;

(2) $P(B)$;

(3) $P(A\cup B)$.

解:

A19. 设 A, B 为两个随机事件, 已知 $P(A)=0.7$, $P(\overline{B})=0.6$, $P(A\overline{B})=0.5$, 求:

(1) $P(A|A\cup B)$;

(2) $P(A|\overline{A}\cup B)$;

(3) $P(AB|A\cup B)$.

解:

A20. 有甲、乙两个盒子, 甲盒中有 3 个红球、2 个白球, 乙盒中有 2 个红球、4 个白球.

(1) 现任选一个盒子, 从中无放回地取 2 球, 求取到的是 1 个红球和 1 个白球的概率;

(2) 从甲盒中取 1 球放入乙盒, 从乙盒中取 1 球, 求从乙盒中取到的球是红球的概率;

(3) 从甲盒中取 1 球放入乙盒, 从乙盒中无放回地取 2 球, 求取到的球是 1 个红球和 1 个白球的概率.

解:

A21. (垃圾邮件过滤) 某人的邮箱收到正常邮件的概率为 0.4, 收到垃圾邮件的概率为 0.6, 正常邮件里包含词语“免费”的概率为 0.005, 垃圾邮件里包含词语“免费”的概率为 0.1. 现在此人设置把含有词语“免费”的邮件自动过滤到垃圾箱中, 求过滤到垃圾箱中的邮件确实是垃圾邮件的概率.

解:

A22. 某公司从甲、乙、丙三家工厂购进同一型号产品, 数量之比为 6:5:4. 已知三家工厂产品的优质品率分别为 85%, 90%, 80%. 若从全部产品中随机取一件, 发现是优质品, 求该产品来自工厂甲、乙、丙的概

率分别是多少, 判断最有可能由哪家工厂生产.

解:

A23. 一架子上有 4 把枪, 其中 3 把是调试好的, 1 把未调试. 设某人用调试好的枪射击时命中率为60%, 用未调试的枪射击时命中率为 5%. 此人从 4 把枪中随机取一把进行射击.

(1) 求命中的概率;

(2) 已知此人命中了, 求取到的是未调试的枪的概率.

解:

A24. 在某一时间点对某证券营业点进行统计. 得知入市时间在 1 年以内的股民赢、平、亏的概率分别为 10%, 20%, 70%; 入市时间在 1 年及以上且不大于 4 年的股民赢、平、亏的概率分别为 20%, 30%, 50%; 入市时间大于 4 年的股民赢、平、亏的概率分别为 50%, 30%, 20%. 入市时间少于 1 年、1 年及以上且不大于 4 年、大于 4 年的股民数分别占 40%, 40%, 20%. 现从该营业点随机找一股民,

(1) 求其有赢利的概率;

(2) 若已知其亏损了,求他为新股民 (入市时间在 1 年以内) 的概率.

解:

A25. 吸烟有害健康是众所周知的事实, 研究结果还发现丈夫每天的吸烟量与胎儿产前的死亡率和先天畸形儿的出生率成正比. 设在父亲不吸烟、每天吸烟 $1\sim10$ 支、每天吸烟 10 支以上条件下, 子女先天畸形的概率分别为 0.8%, 1.4%, 2.1%. 设男性吸烟率为 0.6, 在吸烟的男性中每天吸烟$1\sim10$ 支和每天吸烟 10 支以上的概率分别为 0.3 和 0.7, 并假设父亲的吸烟率与男性的吸烟率相同.

(1) 求先天畸形儿出现的概率;

(2) 若某新生儿出现先天畸形, 求他的父亲每天吸烟 10 支以上的概率.

解:

A26. 某电视台为吸引观众, 对打进电话参与的观众设特等奖一名. 设每个参与观众得奖概率相等. 已知参与观众年龄分布如下表:

年龄 X	$X<20$	$20\leqslant X<40$	$X\geqslant 40$
概率	20%	70%	10%

且各年龄段中女性比例依次为 60%, 50%, 30%. 现知一名女性观众得奖, 求其小于 20 岁的概率.

解:

A27. 证明: 当 $0 < P(B) < 1$ 时, 两事件 A, B 相互独立的充要条件为

$$P(A|B) = P(A|\overline{B}).$$

解:

A28. 设 A, B, C 为三个相互独立的随机事件, 已知 $P(A) = 0.5$, $P(B) = P(C) = 0.4$, 求:

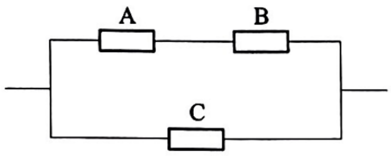
(1) $P(ABC)$;

(2) $P(A \cup B \cup C)$;

(3) $P(AB|AC)$.

解:

A29. 一系统由 3 个独立工作的元件 A, B, C 按下图方式连接而成. 已知元件 A, B, C 正常工作的概率均为 p, 求系统正常工作的概率.



解:

A30. 某人向一目标独立重复射击 3 次, 每次命中率为 0.8, 求:

(1) 前 2 次都命中的概率;

(2) 至少命中 1 次的概率.

解:

A31. 设 A, B, C 为随机事件, $P(A) > 0$, $P(B) > 0$, $P(C) > 0$. 请说明以下说法对还是错:

(1) 若 A, B 相互独立, 则 A, B 相容;

(2) 若 $P(A) = 0.6$, $P(B) = 0.7$, 则 A, B 相互独立;

(3) 若 $P(ABC) = P(A) \cdot P(B) \cdot P(C)$, 则 A, B, C 相互独立;

(4) 若 A, B 互不相容, 则 A, B 不独立.

解:

A32. 连续抛掷一枚硬币, p 表示每次出现正面的概率, 记

$A_i = \{$首次出现正面在第 i 次抛掷$\}$, $i = 1, 2, \cdots$,

$B_j = \{$首次出现连续两次正面在第 $j-1$ 次抛掷和第 j 次抛掷$\}$, $j = 2, 3, \cdots$,

(1) 求 $P(A_i)$ 及 $P(B_4)$, $i = 1, 2, \cdots$;

(2) 求 $P(B_4|A_1)$;

(3) 求 $P(A_1|B_4)$.

解:

A33. 已知一批照明灯管使用寿命大于 1000 h 的概率为 95%, 大于 2000 h 的概率为 30%, 大于 4000 h 的概率为 5%.

(1) 已知一个灯管用了 1000 h 没有坏, 求其使用寿命大于 4000 h 的概率;

(2) 取 10 个灯管独立地装在一大厅内, 过了 2000 h, 求至少有 3 个损坏的概率.

解:

A34. 某玩家独立地向一目标物扔飞镖, 设他命中目标物的概率为 0.05, 问他需要扔多少次才能使至少一次命中目标物的概率超过 0.5?

解:

B1. 设 A, B, C 为三个随机事件, $P(A) = P(B) = 0.3$, $P(C) = 0.4$, 且当 A 发生时 C 必然发生, B 与 C 互不相容, 求:

(1) C 发生的同时 A 不发生的概率;

(2) A 与 B 至少有一个发生的概率;

(3) A, B, C 都不发生的概率.

解:

B2. 在某卫视的一档节目中, 有这样一个项目: 舞台现场摆放了 50 张脸谱, 嘉宾从中随机挑选两张脸谱, 电脑将两张脸谱进行合成产生一张新的脸谱, 挑战者要根据合成的脸谱, 将原先的两张脸谱找出来. 如果靠猜, 求他能够猜中的概率.

解:

B3. 设一社区订 A, B, C 报的家庭分别占 50%, 30%, 40%; 且已知一家庭在订了 A 报的条件下再订 B 报的概率为 20%, 在订了 C 报的条件下再订 A 报或 B 报的概率为 60%. 随机找一家庭, 求该家庭至少订 A, B, C 报中的一种的概率.

解:

B4. 某企业中有 45% 为女职工, 10% 的职工在管理 (技术、质量、行政) 岗位, 5% 的职工为管理岗位的女职工, 在该企业中随机找一位职工.

(1) 已知该职工为女职工, 求该职工在管理岗位的概率;

(2) 已知该职工在管理岗位, 求该职工为男职工的概率.

解:

B5. 某产品 12 个一箱, 成箱出售, 某人购买时随机从中取 2 个检查, 如果没有发现次品就买下. 假设一箱中有 0 个、1 个、2 个次品的概率分别为 0.8, 0.15, 0.05, 求他买下的一箱中的确没有次品的概率.

解:

B6. 有两种同类产品, 第一组有 30 件, 其中 10 件为优质品; 第二组有 20 件, 其中 15 件为优质品. 今从两组中任选一组, 然后从该组中任取 2 次 (每次取 1 件, 无放回抽样).

(1) 求第一次取到的是优质品的概率;

(2) 在已知第 1 次取到的是优质品的条件下, 求第 2 次取到的不是优质品的概率.

解:

B7. 4 个独立元件组成一个系统, 第 i 个元件正常工作的概率为 p_i, $i=1,2,3,4$, 且已知至少有 3 个元件正常工作时系统工作正常.

(1) 求系统正常工作的概率 α;

(2) 在系统正常工作的条件下求 4 个元件均正常工作的概率;

(3) 对系统独立观察 3 次, 求系统恰有 2 次正常工作的概率.

解:

B8. 设某地出现雾霾的概率为 0.4, 在雾霾天, 该地居民独立地以 0.2 的概率戴口罩; 在非雾霾天, 该地居民独立地以 0.01 的概率戴口罩.

(1) 在该地随机选一位居民, 求其戴口罩的概率;

(2) 若在该地同时选 3 位居民, 求至少有一位居民戴口罩的概率.

解:

B9. 某班有 n 个士兵, 每个人各有一支枪. 这些枪在外形上是完全一样的. 在一次夜间紧急集合中每个士兵随机取走一支枪, 求至少有一个士兵拿到自己的枪的概率.

解:

B10. 飞机坠落在 A,B,C 三个区域中一个, 营救部门判断其概率分别为 0.7, 0.2, 0.1. 用直升机搜索这些区域, 若有残骸, 残骸被发现的概率分别为 0.3, 0.4, 0.5. 若已用直升机搜索过 A,B 两个区域, 均未发现残骸, 求在这样的搜救现状下飞机坠落在 C 区域的概率.

解:

B11. 甲、乙、丙三人按照下列规则比赛: 第一局甲、乙参加丙轮空, 优胜者与丙进行第二轮比赛, 失败者轮空, 一直进行到有人连胜两局为止. 连胜两局者称为整场比赛优胜者. 若甲、乙、丙在每局比赛中获胜的概率为 0.5, 求甲、乙、丙成为比赛优胜者的概率分别是多少?

解:

第 2 章　随机变量及其概率分布

A1. 写出下列离散型随机变量 X 的概率分布律:

(1) 从编号为 $1,2,3,4$ 的四个球中, 无放回地取球 2 次, 每次取 1 个球, 令 X 表示取到的两个球中最小的编号;

(2) 将一枚骰子掷 3 次, 令 X 表示出现点数大于 4 的次数;

(3) 一盒子有 2 个红球 3 个白球, 无放回地取球 2 次, 每次取 1 个球, 令 X 表示取到的红球个数;

(4) 一盒子有 2 个红球 3 个白球, 有放回地取球, 每次取 1 个球, 取到红球或取球 3 次就停止取球,令 X 表示取球次数.

解:

(4) X 可能取值为 1,2,3.

A2. 设随机变量 X 的取值为 $1,2,3,4$, 且 $P\{X=i\}=c(5-i)$, $i=1,2,3,4$, 求常数 c 的值及 $P\{1.5<X\leqslant 3\}$.

解:

A3. 设某人一次投篮的命中率是 0.4, 现他独立投篮 3 次, 求:

(1) 他至少投中 2 次的概率;

(2) 他最多投中 2 次的概率.

解:

A4. 设某人独立重复投篮 4 次, 已知他至少投中 1 次的概率为 0.9375, 求他每次投篮的命中率.

解:

A5. 设某人一次投篮的命中率是 0.3, 问他至少要独立投篮多少次, 才能使他至少投中 1 次的概率不小于 0.9?

解:

A6. 设 X 服从参数为 2 的泊松分布,求 $P\{X\leqslant 2\}$, $P\{X\geqslant 2\}$ 及 $P\{X\leqslant 1|X\leqslant 2\}$.

解:

A7. 设某商店某种商品一天的销售量 X (单位:件) 服从参数为 3 的泊松分布,求:

(1) 一天至少售出 4 件的概率;

(2) 一天售出不少于 2 件且不超过 4 件的概率.

解: (1) 所求即为 $P\{X \geqslant 4\}$

(2) 所求即为 $P\{2 \leqslant X \leqslant 4\}$.

A8. 设随机变量 X 的概率分布律为 $P\{X=0\}=0.4$, $P\{X=2\}=0.5$, $P\{X=3\}=0.1$, 求 X 的分布函数及 $P\{1 \leqslant X < 3\}$.

解:

A9. 设随机变量 X 的分布函数

$$F(x)=\begin{cases} 0, & x<-1, \\ 0.3, & -1 \leqslant x<1, \\ 0.7, & 1 \leqslant x<3, \\ 1, & x \geqslant 3, \end{cases}$$

求 X 的概率分布律.

解:

A10. 设随机变量 X 的密度函数为

$$f(x)=\begin{cases} c(2-x), & 0<x<2, \\ 0, & \text{其他}. \end{cases}$$

(1) 求常数 c;

(2) X 的分布函数 $F(x)$;

(3) $P\{0.5<X \leqslant 1\}$.

解:

A11. 设连续型随机变量 X 的分布函数为

$$F(x)=\begin{cases}1-\dfrac{c}{x}, & x\geqslant 2,\\ 0, & x<2.\end{cases}$$

(1) 求常数 c;

(2) 求 X 的密度函数 $f(x)$;

(3) 求 $P\{X\leqslant 4\}$.

解:

A12. 设随机变量 $X\sim U(0,4)$, 求 $\{X^2-2X-3<0\}$ 的概率.

解:

A13. 从区间 $(-1,3)$ 中随机取一个数 X, 写出 X 的密度函数. 若在该区间随机取 $n(n\geqslant 1)$ 个数, 设其中有 Y 个数大于 0, 求 $P\{Y=k\}$, $k=0,1,\cdots,n$.

解:

A14. 设某种产品的寿命 X (单位:年) 服从参数为 0.2 的指数分布.

(1) 求 X 的分布函数;

(2) 求 $P\{X>5\}$;

(3) 求 $P\{X\leqslant 10|X>5\}$.

解:

A15. 设随机变量 $X\sim N(1,4)$.

(1) 求 $P\{X\leqslant 0\}$ 和 $P\{|X-1|\leqslant 2\}$;

(2) 若 $P\{X>a\}=P\{X<a\}$, 求 a;

(3) 求 $P\{|X|\leqslant 2\}$.

解:

A16. 设随机变量 $X\sim N(2,\sigma^2)$. 若要求 $P\{0\leqslant X\leqslant 4\}\geqslant 0.95$, 问 σ 最大取何值?

解:

A17. 设随机变量 X 的概率分布律为 $P\{X=-1\}=0.3$, $P\{X=0\}=0.1$, $P\{X=1\}=0.2$, $P\{X=2\}=0.4$.令 $Y=2X-1$, $Z=X^2$.

(1) 分别求 Y,Z 的概率分布律;

(2) 求 Z 的分布函数.

解:

A18. 设随机变量 X 的密度函数为

$$f(x)=\begin{cases}\dfrac{x}{4}, & 1<x<3,\\ 0, & \text{其他}.\end{cases}$$

在下面三种情况下分别求随机变量 Y 的密度函数:

(1) $Y=2X$;

(2) $Y=2-X$;

(3) $Y=X^2$.

解:

A19. 设随机变量 $X\sim N(1,4)$, 问 $Y=\dfrac{X-1}{2}$ 和 $Z=2-X$ 的概率分布分别是什么？

解:

B1. 从 1,2,3,4,5,6,7 这 7 个数中随机抽取 3 个数(无放回抽样), 并将其从小到大排列, 设排在中间的数为 X, 求 X 的概率分布律.

解:

B2. 某电脑小游戏要依次通过 3 关, 游戏规定过第 1 关和第 2 关各得 1 分, 过第 3 关可得 2 分; 并且规定若其中一关没有通过, 后续关卡仍可进行, 但无论通关与否均不得分. 各个关卡分数累计为游戏总得分. 假设各个关卡的进行是相互独立的(即各个关卡是否通过是相互独立的), 且一玩家通过各关概率均为 20%.

(1) 写出该玩家的游戏总得分 X 的概率分布律;

(2) 求该玩家的游戏总得分大于 2 的微率;

(3) 已知该玩家的游戏总得分不低于 2, 求他得 4 分的概率.

解:

B3. 一个袋中有 6 个球, 其中 3 个是红球, 2 个是白球, 1 个是黑球.从中摸 2 次, 每次摸 1 个球(无放回抽样). 设摸到每一球的概率相等, 记 X 为摸到的红球个数, 写出 X 的概率分布律.

解:

B4. 有人买一种数字型体育彩票,每一注号码中大奖的概率为10^{-7}.

(1) 若每期买 1 注,共买了 n 期, 求他没有中大奖的概率;

(2) 若每期买 10 注 (号码不同), 共买了 n 期, 求他没有中大奖的概率.

解:

B5. 某医院男婴的出生率为 0.51, 如果在该医院随机找 3 名新生儿, 求:

(1) 至少有 1 名男婴的概率;

(2) 恰有 1 名男婴的概率;

(3) 第 1, 第 2 名是男婴, 第 3 名是女婴的概率;

(4) 第 1, 第 2 名是男婴的概率.

解:

B6. 一系统由 5 个独立的同类元件组成, 每个元件正常工作的概率为 0.8, 求:

(1) 恰有 3 个元件正常工作的概率;

(2) 至少有 4 个元件正常工作的概率;

(3) 至多有 2 个元件正常工作的概率.

解:

B7. 一车辆从 A 地到 B 地要经过 3 个特殊地段, 经过这 3 个地段时车辆发生故障的概率分别为 p_1, p_2, p_3. 设在其他地段车辆不发生事故,且记 X 为车辆从 A 地到 B 地发生故障的地段数, Y 为首次发生故障时已通过的特殊地段数 (若没有发生故障,则记 $Y=3$), 分别写出 X 和 Y 的概率分布律.

解:

B8. 从一批不合格率为 $p(0<p<1)$ 的产品中随机抽查产品, 如果查到不合格品就停止检查, 且最多检查 5 件产品. 设停止时已检查了 X 件产品, 求:

(1) X 的概率分布律;

(2) $P\{X \leqslant 2.5\}$.

解:

B9. 设银行自动取款机在单位时间内服务的顾客数 X 服从参数为 1 的泊松分布.

(1) 求单位时间内至少有 2 位顾客接受服务的概率;

(2) 若已知单位时间内至少有 2 位顾客接受服务, 求至多有 3 位顾客接受服务的概率.

解:

B10. 设某地每年生吃鱼胆的人数 X 服从参数为 10 的泊松分布,吃鱼胆而中毒致死的人数 Y 服从参数为 0.5 的泊松分布, 求明年该地:

(1) 至少有 2 人生吃鱼胆的概率;

(2) 没有人因生吃鱼胆致死的概率.

解:

B11. 某公交车站单位时间内候车人数服从参数为 λ 的泊松分布.

(1) 若已知单位时间内至少有 1 人候车的概率为 $(1-e^{-4.5})$, 求单位时间内至少有 2 人候车的概率;

(2) 若 $\lambda=3.2$, 且已知至少有 1 人在此候车, 求该车站只有他 1 人候车的概率.

解:

B12. 设某手机在早上 9:00 至晚上 9:00 的任意长度为 t (单位: min) 的时间区间内收到的短信数 X 服从参数为 λt 的泊松分布, $\lambda=\dfrac{1}{20}$, 且与时间起点无关.

(1) 求 10:00 到 12:00 期间恰好收到 6 条短信的概率;

(2) 已知在 10:00 到 12:00 期间至少收到 5 条短信, 求在该时段恰好收到 6 条短信的概率.

解:

B13. 某大学每年 5 月份组织教职工体检: 根据以往的情况, 通过体检发现千分之一的被检者患有重大疾病. 已知有 3000 人参加今年的体检, 求至少有 2 人被检出重大疾病的概率的近似值 (用泊松分布来近似计算).

解:

B14. 一袋中有 10 个球,编号为 $0,1,\cdots,9$.

(1) 采用无放回抽样取 3 次, 每次取 1 球, X 表示所取球的号码大于 6 的个数, 求 X 的概率分布律;

(2) 采用有放回抽样取 3 次, 每次取 1 球, Y 表示所取球的号码为偶数的个数, 求 Y 的概率分布律;

(3) 采用有放回抽样取球, 直到取到号码 9 为止, Z 表示取球次数, 求 Z 的概率分布律;

(4) 采用有放回抽样取球, 求第 5 次恰好取到第 3 个奇数号码球的概率.

解:

B15. 小王租到一所房子, 房东给了他 5 把钥匙, 其中只有一把能打开大门. 计算在以下两种方式下, 他打开大门所需的试钥匙次数的概率分布律:

(1) 每次都从全部 5 把钥匙中任选一把试开;

(2) 每次试开失败后, 将该把钥匙单独放置, 从剩余的钥匙中任选一把试开.

解:

B16. 设随机变量 X 具有以下性质:

当 $0 \leqslant X \leqslant 1$ 时, $P\{0 \leqslant X \leqslant 1\} = \dfrac{x}{2}$;

当 $2 \leqslant X \leqslant 3$ 时, $P\{2 \leqslant X \leqslant 3\} = \dfrac{x-2}{2}$.

(1) 写出 X 的分布函数;

(2) 求 $P\{X \leqslant 2.5\}$.

解:

B17. 设随机变量 X 的密度函数为

$$f(x) = \begin{cases} c(4-x^2), & 0 < x < 2, \\ 0, & \text{其他}. \end{cases}$$

(1) 求常数 c;

(2) 求 X 的分布函数 $F(x)$;

(3) 求 $P\{-1 < X < 1\}$;

(4) 对 X 独立观察 5 次, 求事件 $\{-1 < X < 1\}$ 恰好发生 2 次的概率.

解:

B18. 设连续型随机变量 X 的分布函数为

$$F(x) = \begin{cases} 0, & x < 0, \\ ax^2, & 0 \leqslant x < 1, \\ bx, & 1 \leqslant x < 2, \\ 1, & x \geqslant 2. \end{cases}$$

(1) 求常数 a, b;

(2) 求 X 的密度函数 $f(x)$;

(3) 求 $P\{0.5 < X < 1.5\}$.

解:

B19. 已知在早上 $7:00 - 8:00$ 有两班车从 A 校区到 B 校区, 出发时间分别是 7:30 和 7:50, 一学生在 $7:20 - 7:45$ 随机到达车站乘这两班车.

(1) 求该学生等车时间小于 10 min 的概率;

(2) 求该学生等车时间大于 5 min 且小于 15 min 的概率;

(3) 已知该学生等车时间大于 5 min 的条件下, 求他能赶上 7:30 这班车的概率.

解:

B20. 设随机变量 X 服从正态分布 $N(\mu, \sigma^2)$, 其中 $\mu = 5$, $\sigma = 1$.

(1) 求 $P\{X > 2.5\}$;

(2) 求 $P\{X < 3.52\}$;

(3) 求 $P\{4 < X < 6\}$;

(4) 求 $P\{|X - 5| > 2\}$;

解:

B21. 设某人的年收入扣除日常花费后的余额(单位: 万元) X 服从正态分布 $N(6.5,1)$, 且往年没有积蓄, 也不打算借贷, 今年他计划至少花 7 万元买家电, 求他能实现自己计划的概率.

解:

B22. 设一地区的青年男子身高 (单位: cm) X 服从正态分布 $N(170,5.0^2)$. 现在该地区随机找一青年男子测身高, 求:

(1) 身高大于 170 cm 的概率;

(2) 身高大于 165 cm 且小于 175 cm 的概率;

(3) 身高小于 172 cm 的概率.

解:

B23. 设某群体的 BMI (体重指数)值 (单位: kg/m^2) $X \sim N(22.5,2.5^2)$. 医学研究发现身体肥胖者患高血压的可能性增大: 当 $X \leqslant 25$ 时, 患高血压的概率为 10%; 当 $25 < X \leqslant 27.5$ 时, 患高血压的概率为 15%; 当 $X > 27.5$ 时, 患高血压的概率为 30%.

(1) 从该群体中随机选出 1 人,求他患高血压的概率;

(2) 若他患有高血压,求他的 BMI 值超过 25 的概率;

(3)随机独立地选出 3 人, 求至少有 1 人患高血压的概率.

解:

B24. 设系统电压 (单位: V) 在小于 200, 在区间 $[200,240]$ 上和超过 240 这三种情况下, 系统中某种电子元件不能正常工作的概率分别为 0.1,0.001,0.2. 设系统电压 X 服从 $N(220,25^2)$.

(1) 求该电子元件不能正常工作的概率 α;

(2) 若该电子元件不能正常工作, 求此时系统电压超过 240 V 的概率 β;

(3) 设某系统有 3 个这种元件, 且若至少有 2 个正常时系统才运行正常, 求该系统运行正常的概率 θ.

解:

B25. 设一高速公路某处双休日一天车流量 $X \sim N(\mu,\sigma^2)$, 有 30% 的天数车流量小于 12800 辆, 有 95% 的天数车流量大于 10000辆, 求 μ, σ.

解:

B26. 设随机变量 X 服从 $N(15,4)$, X 落在区间 $(-\infty,x_1]$, $(x_1,x_2]$, $(x_2,+\infty)$ 中的概率之比为 50:34:16, 求 x_1,x_2 的值.

解:

B27. 设随机变量 X 的密度函数为

$$f(x) = a \cdot e^{-x^2}, \quad -\infty < x < \infty.$$

(1) 求常数 a;

(2) 求 $P\left\{X > \frac{1}{2}\right\}$.

解:

B28. 设银行某一柜台一位顾客的服务时间 (单位: min) 服从参数为 $\lambda = \frac{1}{8}$ 的指数分布. 若在顾客 A 到达时恰好有 1 人正在接受服务, 且无其他人排队, 设 A 的等待时间为 X.

(1) 求 X 的密度函数;

(2) 求 A 等待时间超过 10 min 的概率;

(3) 求等待时间大于 8 min 且小于 16 min 的概率.

解:

B29. 设甲、乙两厂生产的同类型产品寿命 (单位: 年) 分别服从参数为 $\frac{1}{3}$ 和 $\frac{1}{6}$ 的指数分布, 将两厂的产品混在一起, 其中甲厂的产品数占 40%. 现从这批混合产品中随机取一件.

(1) 求该产品寿命大于 6 年的概率;

(2) 若已知取到的是甲厂产品, 在已用了 4 个月没有坏的条件下, 求其用到 1 年还不坏的概率;

(3) 在该产品已用了 4 个月没有坏的条件下, 求其用到 1 年还不坏的概率.

解:

B30. 以 X 表示某商店早晨开门后直到第一个顾客到达的等待时间 (单位: min), X 的分布函数为

$$F(x) = \begin{cases} 1 - e^{-0.2x}, & x \geqslant 0, \\ 0, & x < 0. \end{cases}$$

(1) 求 X 的密度函数 $f(x)$;

(2) 求 $P\{5 < X < 10\}$;

(3) 求某一周 (7 天) 至少有 6 天等待时间不超过 5 min 的概率.

解:

B31. 设一批电子元件寿命 X (单位: h) 的密度函数为

$$f(x)=\begin{cases}0.01e^{-0.01x}, & x>0,\\ 0, & x\le 0.\end{cases}$$

某人买了 3 个元件试用, 若至少有 2 个寿命大于 150 h, 则下次再买此类元件.

(1) 求这 3 个元件中恰好有 2 个寿命大于 150h 的概率;

(2) 求这个人会再买的概率.

解:

B32. 某次游戏向每个玩家发 5 个球, 向目标投掷, 投中 2 次就结束投球. 若每次投中的概率均为 $p=0.7$, 且每次投掷是相互独立的. 设 X 为结束时的投球次数, 规定当 $X=2$ 时得 10 分, 当 $X=3$ 时得 8 分, 当 $X\geqslant 4$ 时得 2 分, 记 Y 为所得分数, 写出 Y 的概率分布律.

解:

B33. 已知随机变量 X 的密度函数为

$$f(x)=\begin{cases}c(4-x^2), & -1<x<2,\\ 0, & \text{其他}.\end{cases}$$

(1) 求常数 c;

(2) 设 $Y=3X$, 求 Y 的密度函数;

(3) 设 $Z=|X|$, 求 Z 的分布函数及密度函数.

解:

B34. 设在时间区间 $(0,t]$ 内进入某商店的顾客数 $N(t)$ 服从参数为 λt 的泊松分布,且设第 1 个顾客到达时间为 T.

(1) 求 T 的分布函数;

(2) 求 $P\{T>t_0+t|T>t_0\}$, 其中 $t>0,t_0>0$.

解:

B35. 从区间 $(0,1)$ 上随机取一数 X, 记 $Y=X^n$ $(n>1,n$ 为整数), 求 Y 的密度函数.

解:

B36. 设随机变量 X 服从 $\left(0, \dfrac{3\pi}{2}\right)$ 上的均匀分布, $Y = \cos X$, 求 Y 的分布函数.

解:

B37. 设随机变量 $X \sim N(\mu, \sigma^2)$, 求 $Y = X^2$ 的密度函数.

解:

B38. 设随机变量 X 的密度函数为

$$f(x) = \begin{cases} ax + b, & 0 < x < 2, \\ 0, & \text{其他}. \end{cases}$$

且已知 $P\{X < 1\} = \dfrac{1}{3}$.

(1) 求常数 a, b;

(2) 设 $Y = \sqrt{X}$, 求 Y 的密度函数 $f_Y(y)$.

解:

B39. 设随机变量 $X \sim N(0, 1)$, 记 $Y = e^X$, $Z = \ln|X|$.

(1) 求 Y 的密度函数;

(2) 求 Z 的密度函数.

解:

第 3 章　多维随机变量及其分布

A1. 设二维离散型随机变量(X,Y) 的联合分布律为

X \ Y	0	1	2	3
0	0.1	0.1	$2c$	0.1
1	0	0.1	0.1	0
2	c	0	0	0.2

(1) 求常数 c;

(2) 求 $P\{X \leqslant 1, Y \geqslant 1\}$;

(3) 分别求 X 和 Y 的边际分布律.

解: (1) 由联合分布律, $0.1+0.1+2c+0.1+0.1+0.1+c+0.2=3c+0.7=1$, 则 $c=0.1$.

(2) $P\{X \leqslant 1, Y \geqslant 1\}=0.1+2c+0.1+0.1+0.1+0=0.6$.

(3) 对 X, 其边际分布律为 $P\{X=0\}=0.5$, $P\{X=1\}=0.2$, $P\{X=2\}=0.3$.

Y 的边际分布律为 $P\{Y=0\}=0.2$, $P\{Y=1\}=0.2$, $P\{Y=2\}=0.3$, $P\{Y=3\}=0.3$.

A2. 设二维离散型随机变量 (X,Y) 的联合分布律为

X \ Y	0	1	2
0	0.1	a	0.1
1	0	0.1	$2a$
2	b	0	0.2

分别在下列条件下求常数 a 和 b:

(1) $P\{X \leqslant 1\}=0.6$;

(2) $P\{X=0|Y=0\}=0.5$;

(3) $P\{X \leqslant 1, Y \geqslant 1\}=0.35$.

解: 由联合分布律得 $0.1+a+0.1+0.1+2a+b+0.2=1$, 即 $3a+b=0.5$.

(1) $P\{X \leqslant 1\}=0.1+a+0.1+0.1+2a=0.3+3a=0.6$.

则 $a=0.1$, $b=0.2$.

(2) $P\{X=0|Y=0\}=\dfrac{P\{X=0,Y=0\}}{P\{Y=0\}}=\dfrac{0.1}{0.1+b}=0.5$.

则 $b=0.1$, $a=\dfrac{0.4}{3}=\dfrac{2}{15}$.

(3) $P\{X\leqslant 1,Y\geqslant 1\}=a+0.1+0.1+2a=0.35$.

则 $a=0.05$, $b=0.35$.

A3. 有两个袋子均放着 3 个红球 2 个白球, 今从两个袋子中同时各摸出 1 个球互换 (设每个袋子摸到每个球的概率相等). 记 X,Y 分别为两个袋子中互换球后的红球个数, 求 (X,Y) 的联合分布律及 X 的边际分布律.

解: 当两个袋子同时摸出红球或同时摸出白球时, 互换后两个袋子球的情况不变.

即此时 $X=Y=3$. 故 $P\{X=3,Y=3\}=\dfrac{3}{5}\cdot\dfrac{3}{5}+\dfrac{2}{5}\cdot\dfrac{2}{5}=\dfrac{13}{25}=0.52$.【都抽红+都抽白】

当 X 对应袋子抽红, Y 对应袋子抽白时, 交换后 $X=2$, $Y=4$.

此时 $P\{X=2,Y=4\}=\dfrac{3}{5}\cdot\dfrac{2}{5}=\dfrac{6}{25}=0.24$.

当 X 对应袋子抽白, Y 对应袋子抽红时, 交换后 $X=4$, $Y=2$.

此时 $P\{X=4,Y=2\}=\dfrac{2}{5}\cdot\dfrac{3}{5}=\dfrac{6}{25}=0.24$.

从而 (X,Y) 的联合分布律为

X	Y		
	2	3	4
2	0	0	0.24
3	0	0.52	0
4	0.24	0	0

X 的边际分布律 $P\{X=2\}=0.24$, $P\{X=3\}=0.52$, $P\{X=4\}=0.24$.

A4. 盒子中有 3 个红球和 2 个白球, 取 2 次球, 每次取 1 个. 设

$$X=\begin{cases}1, & \text{若第 1 次取到红球,}\\ 0, & \text{若第 1 次取到白球.}\end{cases}$$

$$Y=\begin{cases}1, & \text{若第 2 次取到红球,}\\ 0, & \text{若第 2 次取到白球.}\end{cases}$$

分别求在无放回抽样和有放回抽样这两种情况下 (X,Y) 的联合分布律及 X 和 Y 的边际分布律.

解: (1) 无放回抽样时, 第一次抽取结果影响第二次的抽取结果.

$P\{X=0,Y=0\}=\frac{2}{5}\times\frac{1}{4}=\frac{1}{10}=0.1;\qquad P\{X=0,Y=1\}=\frac{2}{5}\times\frac{3}{4}=\frac{3}{10}=0.3.$

$P\{X=1,Y=0\}=\frac{3}{5}\times\frac{2}{4}=\frac{3}{10}=0.3;\qquad P\{X=1,Y=1\}=\frac{3}{5}\times\frac{2}{4}=\frac{3}{10}=0.3.$

从而 (X,Y) 联合分布律为

X \ Y	0	1
0	0.1	0.3
1	0.3	0.3

此时 $P\{X=0\}=0.4$, $P\{X=1\}=0.6$.

$P\{Y=0\}=0.4$, $P\{Y=1\}=0.6$.

(2) 有放回时, 每次取到白球概率为 $\frac{2}{5}$, 取到红球概率为 $\frac{3}{5}$, 且此时两次取球独立.

$P\{X=0,Y=0\}=P\{X=0\}\cdot P\{Y=0\}=\frac{2}{5}\times\frac{2}{5}=\frac{4}{25}=0.16.$

$P\{X=0,Y=1\}=P\{X=0\}\cdot P\{Y=1\}=\frac{2}{5}\times\frac{3}{5}=\frac{6}{25}=0.24.$

$P\{X=1,Y=0\}=P\{X=1\}\cdot P\{Y=0\}=\frac{3}{5}\times\frac{2}{5}=\frac{6}{25}=0.24.$

$P\{X=1,Y=1\}=P\{X=1\}\cdot P\{Y=1\}=\frac{3}{5}\times\frac{3}{5}=\frac{9}{25}=0.36.$

从而 (X,Y) 联合分布律为

X \ Y	0	1
0	0.16	0.24
1	0.24	0.36

此时 $P\{X=0\}=0.4$, $P\{X=1\}=0.6$.

$P\{Y=0\}=0.4$, $P\{Y=1\}=0.6$.

A5. 设二维离散型随机变量 (X,Y) 的联合分布律为

X \ Y	0	1
0	0.3	a
1	b	0.2

且已知事件 $\{X=0\}$ 与事件 $\{X+Y=1\}$ 相互独立, 求常数 a 和 b 的值.

解: 由联合分布律, $0.3+a+b+0.2=1$, 即 $a+b=0.5$.

又因为 $P\{X+Y=1\}=P\{X=0,Y=1\}+P\{X=1,Y=0\}=a+b=0.5$,

$P\{X=0\}=0.3+a$, $P\{X=0,X+Y=1\}=P\{X=0,Y=1\}=a$.

由事件独立有 $P\{X=0,X+Y=1\}=P\{X=0\}\cdot P\{X+Y=1\}$.

即 $a=0.5(0.3+a)$, 解得 $a=0.3$, 从而 $b=0.2$.

A6. 设二维离散型随机变量 (X,Y) 的联合分布律为

X \ Y	-1	0	1
1	a	0.1	b
2	0.1	0.1	c

且已知 $P\{Y\leqslant 0|X<2\}=0.5$, $P\{Y=1\}=0.5$, 求 a, b, c 的值及 X 和 Y 的边际分布律.

解: 由题, $a+0.1+b+0.1+0.1+c=1$, 且 $P\{Y=1\}=b+c=0.5$.

从而有 $a=0.2$, 则 $P\{Y\leqslant 0,X<2\}=a+0.1=0.3$, $P\{X<2\}=P\{X=1\}=a+0.1+b=0.3+b$.

因此 $P\{Y\leqslant 0|X<2\}=\dfrac{P\{Y\leqslant 0,X<2\}}{P\{X<2\}}=\dfrac{0.3}{0.3+b}=0.5$, 则 $b=0.3$, 因此 $c=0.2$.

从而对于 X, $P\{X=1\}=0.6$, $P\{X=2\}=0.4$.

对于 Y, $P\{Y=-1\}=0.3$, $P\{Y=0\}=0.2$, $P\{Y=1\}=0.5$.

A7. 设二维离散型随机变量 (X,Y) 的联合分布律为

X \ Y	0	1	2
0	0.1	0.1	0
1	0	0.2	0.2
2	0.2	0	0.2

(1) 求给定 $\{X=1\}$ 的条件下 Y 的条件分布律;

(2) 求给定 $\{Y=1\}$ 的条件下 X 的条件分布律.

解: (1) $X=1$ 时 Y 可以取 $1,2$.

因为 $P\{X=1\}=0.2+0.2=0.4$, $P\{X=1,Y=1\}=P\{X=1,Y=2\}=0.2$.

条件分布律为

$$P\{Y=1|X=1\}=\frac{P\{X=1,Y=1\}}{P\{X=1\}}=0.5,\quad P\{Y=2|X=1\}=\frac{P\{X=1,Y=2\}}{P\{X=1\}}=0.5.$$

(2) $Y=1$ 时 X 可以取 $0,1$.

因为 $P\{Y=1\}=0.1+0.2=0.3$, $P\{X=0,Y=1\}=0.1$, $P\{X=1,Y=1\}=0.2$.

条件分布律为

$$P\{X=0|Y=1\}=\frac{P\{X=0,Y=1\}}{P\{Y=1\}}=\frac{1}{3},\quad P\{X=1|Y=1\}=\frac{P\{X=1,Y=1\}}{P\{Y=1\}}=\frac{2}{3}.$$

A8. 设随机变量 X,Y 的概率分布律分别为

X	0	1
p	0.4	0.6

Y	0	1	2
p	0.2	0.5	0.3

且已知 $P\{X=0,Y=0\}=P\{X=1,Y=2\}=0.2$.

(1) 写出 (X,Y) 的联合分布律;

(2) 写出给定 $\{X=0\}$ 的条件下 Y 的条件分布律.

解: (1) 因为 $P\{Y=2\}=0.3$, 则 $P\{X=0,Y=2\}=P\{Y=2\}-P\{X=1,Y=2\}=0.1$.

又因为 $P\{Y=0\}=0.2$, 则 $P\{X=1,Y=0\}=P\{Y=0\}-P\{X=0,Y=0\}=0$.

则 $P\{X=1,Y=1\}=P\{X=1\}-P\{X=1,Y=0\}-P\{X=1,Y=2\}=0.4$.

$P\{X=0,Y=1\}=P\{Y=1\}-P\{X=1,Y=1\}=0.1$.

从而 (X,Y) 的联合分布律为

X	Y		
	0	1	2
0	0.2	0.1	0.1
1	0	0.4	0.2

(2) $P\{Y=0|X=0\}=\dfrac{P\{X=0,Y=0\}}{P\{X=0\}}=\dfrac{0.2}{0.4}=0.5$;

$P\{Y=1|X=0\}=\dfrac{P\{X=0,Y=1\}}{P\{X=0\}}=\dfrac{0.1}{0.4}=0.25$;

$P\{Y=2|X=0\}=\dfrac{P\{X=0,Y=2\}}{P\{X=0\}}=\dfrac{0.1}{0.4}=0.25$.

A9. 将一枚均匀的骰子抛 2 次, 记 X 为第 1 次出现的点数, Y 为 2 次点数的最大值.

(1) 求 (X,Y) 的联合分布律及边际分布律;

(2) 写出给定 $\{Y=6\}$ 的条件下 X 的条件分布律.

解: (1) 此时记第二次出现点数为 Z, 则 X, Y, Z 可能取值均为 1,2,3,4,5,6.

当 $Y=1$ 时, 则 $X=1$, 此时两次都是 1 点. $P\{X=1,Y=1\}=P\{X=1,Z=1\}=\frac{1}{6}\times\frac{1}{6}=\frac{1}{36}$.

当 $Y=2$ 时, X 只能取 1 或 2. 此时有

$$P\{X=1,Y=2\}=P\{X=1,Z=2\}=\frac{1}{6}\times\frac{1}{6}=\frac{1}{36}.$$

$$P\{X=2,Y=2\}=P\{X=2,Z\leqslant 2\}=\frac{1}{6}\times\frac{2}{6}=\frac{2}{36}.$$

当 $Y=3$ 时, X 取 1,2,3. 此时有

$$P\{X=1,Y=3\}=P\{X=1,Z=3\}=\frac{1}{6}\times\frac{1}{6}=\frac{1}{36}.$$

$$P\{X=2,Y=3\}=P\{X=2,Z=3\}=\frac{1}{6}\times\frac{1}{6}=\frac{1}{36}.$$

$$P\{X=3,Y=3\}=P\{X=3,Z\leqslant 3\}=\frac{1}{6}\times\frac{3}{6}=\frac{3}{36}.$$

当 $Y=4$ 时, X 取 1,2,3,4. 此时有

$$P\{X=1,Y=4\}=P\{X=1,Z=4\}=\frac{1}{6}\times\frac{1}{6}=\frac{1}{36}.$$

$$P\{X=2,Y=4\}=P\{X=2,Z=4\}=\frac{1}{6}\times\frac{1}{6}=\frac{1}{36}.$$

$$P\{X=3,Y=4\}=P\{X=3,Z=4\}=\frac{1}{6}\times\frac{1}{6}=\frac{1}{36}.$$

$$P\{X=4,Y=4\}=P\{X=4,Z\leqslant 4\}=\frac{1}{6}\times\frac{4}{6}=\frac{4}{36}.$$

当 $Y=5$ 时, X 取 1,2,3,4,5. 此时有

$$P\{X=1,Y=5\}=P\{X=1,Z=5\}=\frac{1}{6}\times\frac{1}{6}=\frac{1}{36}.$$

$$P\{X=2,Y=5\}=P\{X=2,Z=5\}=\frac{1}{6}\times\frac{1}{6}=\frac{1}{36}.$$

$$P\{X=3,Y=5\}=P\{X=3,Z=5\}=\frac{1}{6}\times\frac{1}{6}=\frac{1}{36}.$$

$$P\{X=4,Y=5\}=P\{X=4,Z=5\}=\frac{1}{6}\times\frac{1}{6}=\frac{1}{36}.$$

$$P\{X=5,Y=5\}=P\{X=5,Z\leqslant 5\}=\frac{1}{6}\times\frac{5}{6}=\frac{5}{36}.$$

当 $Y=6$ 时, X 取 1,2,3,4,5,6. 且 $X=6$ 时 Y 一定为 6. 此时有

$$P\{X=1,Y=6\}=P\{X=1,Z=6\}=\frac{1}{6}\times\frac{1}{6}=\frac{1}{36}.$$

$$P\{X=2,Y=6\}=P\{X=2,Z=6\}=\frac{1}{6}\times\frac{1}{6}=\frac{1}{36}.$$

$$P\{X=3,Y=6\}=P\{X=3,Z=6\}=\frac{1}{6}\times\frac{1}{6}=\frac{1}{36}.$$

$$P\{X=4,Y=6\}=P\{X=4,Z=6\}=\frac{1}{6}\times\frac{1}{6}=\frac{1}{36}.$$

$$P\{X=5,Y=6\}=P\{X=5,Z=6\}=\frac{1}{6}\times\frac{1}{6}=\frac{1}{36}.$$

$$P\{X=6,Y=6\}=P\{X=6\}=\frac{1}{6}=\frac{6}{36}.$$

从而 (X,Y) 的联合分布律为

X \ Y	1	2	3	4	5	6
1	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
2	0	$\frac{2}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
3	0	0	$\frac{3}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
4	0	0	0	$\frac{4}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
5	0	0	0	0	$\frac{5}{36}$	$\frac{1}{36}$
6	0	0	0	0	0	$\frac{6}{36}$

因此边界分布律为

X	1	2	3	4	5	6
p	$\frac{6}{36}$	$\frac{6}{36}$	$\frac{6}{36}$	$\frac{6}{36}$	$\frac{6}{36}$	$\frac{6}{36}$

Y	1	2	3	4	5	6
p	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	$\frac{11}{36}$

(2) 由题, $P\{X=1|Y=6\}=\dfrac{P\{X=1,Y=6\}}{P\{Y=6\}}=\dfrac{\frac{1}{36}}{\frac{11}{36}}=\dfrac{1}{11}$,

$$P\{X=2|Y=6\}=\frac{P\{X=2,Y=6\}}{P\{Y=6\}}=\frac{\frac{1}{36}}{\frac{11}{36}}=\frac{1}{11},$$

$$P\{X=3|Y=6\}=\frac{P\{X=3,Y=6\}}{P\{Y=6\}}=\frac{\frac{1}{36}}{\frac{11}{36}}=\frac{1}{11},$$

$$P\{X=4|Y=6\}=\frac{P\{X=4,Y=6\}}{P\{Y=6\}}=\frac{\frac{1}{36}}{\frac{11}{36}}=\frac{1}{11},$$

$$P\{X=5|Y=6\}=\frac{P\{X=5,Y=6\}}{P\{Y=6\}}=\frac{\frac{1}{36}}{\frac{11}{36}}=\frac{1}{11},$$

$$P\{X=6|Y=6\}=\frac{P\{X=6,Y=6\}}{P\{Y=6\}}=\frac{\frac{6}{36}}{\frac{11}{36}}=\frac{6}{11}.$$

A10. 设一大型设备单位时间内发生的故障数 X 具有概率分布律

X	0	1	2
p	0.6	0.3	0.1

每次故障以概率 p 带来损失 a 万元. 设 Y 为该设备在单位时间内间内的损失 (单位: 万元).

(1) 求 (X,Y) 的联合分布律;

(2) 在该设备仅发生 1 次故障的条件下, 求 Y 的条件分布律.

解: (1) $X=0$ 时, $Y=0$, 此时 $P\{X=0,Y=0\}=P\{X=0\}=0.6$.

$X=1$ 时, Y 可以取 0, a.

$P\{X=1,Y=0\}=P\{X=1\}(1-p)=0.3(1-p)$, $P\{X=1,Y=a\}=P\{X=1\}\cdot p=0.3p$.

$X=2$ 时, Y 可以取 0, a, $2a$.

$P\{X=2,Y=0\}=P\{X=2\}(1-p)^2=0.1(1-p)^2$,

$P\{X=2,Y=a\}=P\{X=2\}\cdot 2p(1-p)=0.2p(1-p)$,

$P\{X=2,Y=2a\}=P\{X=2\}\cdot p^2=0.1p^2$.

从而 (X,Y) 的联合分布律为

X	Y		
	0	a	$2a$
0	0.6	0	0
1	$0.3(1-p)$	$0.3p$	0
2	$0.1(1-p)^2$	$0.2p(1-p)$	$0.1p^2$

(2) 仅发生一次故障即 $X=1$.

$$P\{Y=0|X=1\}=\frac{P\{Y=0,X=1\}}{P\{X=1\}}=\frac{0.3(1-p)}{0.3}=1-p.$$

$$P\{Y=a|X=1\}=\frac{P\{Y=a,X=1\}}{P\{X=1\}}=\frac{0.3p}{0.3}=p.$$

A11. 设二维离散型随机变量 (X,Y) 的联合分布律为

X	Y		
	0	1	2
0	0.1	0.1	0
1	0	0.2	0.2
2	0.2	0	0.2

记$F(x,y)$, $F_X(x)$ 分别为 (X,Y) 的联合分布函数和 X 的边际分布函数.

(1) 求 $F(0,1)$, $F(1,1.5)$, $F(2.1,1.1)$;

(2) 求 $F_X(x)$.

解: (1) $F(0,1)=P\{X\leqslant 0,Y\leqslant 1\}=0.1+0.1=0.2$.

$F(1,1.5)=P\{X\leqslant 1,Y\leqslant 1.5\}=0.1+0+0.1+0.2=0.4$.

$F(2.1,1.1)=P\{X\leqslant 2.1,Y\leqslant 1.1\}=0.1+0+0.2+0.1+0.2+0=0.6$.

(2) $x<0$ 时, x 不可能大于 X 的取值, 则 $F_X(x)=P\{X\leqslant x\}=0$;

$0\leqslant x<1$ 时, $F_X(x)=P\{X\leqslant x\}=P\{X=0\}=0.1+0.1=0.2$;

$1\leqslant x<2$ 时, $F_X(x)=P\{X\leqslant x\}=P\{X=0\}+P\{X=1\}=0.2+0.2+0.2=0.6$;

$x\geqslant 2$ 时, 则 $X\leqslant x$ 恒成立, 有 $F_X(x)=P\{X\leqslant x\}=1$.

综上, $F_X(x)=\begin{cases}0, & x<0,\\ 0.2, & 0\leqslant x<1,\\ 0.6, & 1\leqslant x<2,\\ 1, & x\geqslant 2.\end{cases}$

A12. 设二维离散型随机变量 (X,Y) 的边际分布函数

$$F_X(x)=\begin{cases}0, & x<1,\\ 0.3, & 1\leqslant x<2,\\ 1, & x\geqslant 2,\end{cases}\qquad F_Y(y)=\begin{cases}0, & y<0,\\ 0.4, & 0\leqslant x<1,\\ 1, & x\geqslant 1,\end{cases}$$

且已知 $P\{X=1,Y=0\}=0.1$.

(1) 求 (X,Y) 的联合分布律;

(2) 求给定 $\{Y=0\}$ 的条件下 X 的条件分布函数.

解: 由分布函数知, X 取值为 1,2; Y 取值为 0,1. 且 $P\{X=1\}=0.3$, $P\{Y=0\}=0.4$.

因此 $P\{X=2\}=1-0.3=0.7$, $P\{Y=1\}=1-0.4=0.6$. 从而

$$P\{X=1,Y=1\}=P\{X=1\}-P\{X=1,Y=0\}=0.2,$$

$$P\{X=2,Y=0\}=P\{Y=0\}-P\{X=1,Y=0\}=0.3,$$

$$P\{X=2,Y=1\}=P\{X=2\}-P\{X=2,Y=0\}=0.4.$$

从而 (X,Y) 的联合分布律为

X \ Y	0	1
1	0.1	0.2
2	0.3	0.4

(2) 给定 $Y=0$ 时, 有 $P\{X=1|Y=0\}=\dfrac{P\{X=1,Y=0\}}{P\{Y=0\}}=\dfrac{0.1}{0.4}=0.25$.

因为 X 只能取 1 或 2, 则对 $x<1$, $P\{X\leqslant x|Y=0\}=0$; 对 $x\geqslant 2$, $P\{X\leqslant x|Y=0\}=1$.

对 $1\leqslant x<2$, $P\{X\leqslant x|Y=0\}=P\{X=1|Y=0\}=0.25$.

因此条件分布函数 $F_{X|Y}(x|0)=\begin{cases}0, & x<1,\\ 0.25, & 1\leqslant x<2,\\ 1, & x\geqslant 2.\end{cases}$

A13. 设 A,B 为两随机事件, 已知 $P(A)=0.3$, $P(B|\overline{A})=0.5$, $P(B)=0.4$. 引入随机变量 X,Y, 分别为

$$X=\begin{cases}1, & A\text{发生},\\ 0, & A\text{不发生},\end{cases}\qquad Y=\begin{cases}1, & B\text{发生},\\ 0, & B\text{不发生}.\end{cases}$$

(1) 求 (X,Y) 的联合分布律;

(2) 求 X 的边际分布函数;

(3) 求给定 $\{X=1\}$ 的条件下 Y 的条件分布函数.

解: 由题, $P(\overline{A})=0.7$. 则 $P(\overline{A}B)=P(\overline{A})P(B|\overline{A})=0.35$.

则 $P\{X=1,Y=1\}=P(AB)=P(B)-P(\overline{A}B)=0.05$,

$P\{X=1,Y=0\}=P(A\overline{B})=P(A)-P(AB)=0.25$,

$P\{X=0,Y=1\}=P(\overline{A}B)=0.35$, $P\{X=0,Y=0\}=P(\overline{AB})=P(\overline{A})-P(\overline{A}B)=0.35$.

从而 (X,Y) 的联合分布律为

X	Y	
	0	1
0	0.35	0.35
1	0.25	0.05

(2) 因为 X 只能取 0 或 1, 则 $x<0$ 时 $P\{X\leqslant x\}=0$; $x\geqslant 1$ 时 $P\{X\leqslant x\}=1$.

又因为 $P\{X=0\}=P(\overline{A})=0.7$, 则 $0\leqslant x<1$ 时 $P\{X\leqslant x\}=0.7$.

因此 X 的边界分布函数 $F_X(x)=\begin{cases}0, & x<0,\\ 0.7, & 0\leqslant x<1,\\ 1, & x\geqslant 1.\end{cases}$

(3) 给定 $X=1$ 时, 因为 Y 只能取 0 或 1,

对 $y<0$, $P\{Y\leqslant y|X=1\}=0$; 对 $y\geqslant 1$, $P\{Y\leqslant y|X=1\}=1$.

又因为 $P\{Y=0|X=1\}=\dfrac{P\{Y=0,X=1\}}{P\{X=1\}}=\dfrac{0.25}{0.3}=\dfrac{5}{6}$.

对 $0\leqslant y<1$, $P\{Y\leqslant y|X=1\}=P\{Y=0|X=1\}=\dfrac{5}{6}$.

因此条件分布函数 $F_{Y|X}(y|1)=\begin{cases}0, & y<0,\\ \dfrac{5}{6}, & 0\leqslant y<1,\\ 1, & y\geqslant 1.\end{cases}$

A14. 设二维连续型随机变量 (X,Y) 的联合密度函数为

$$f(x,y)=\begin{cases}c+xy, & 0<x<1,0<y<1,\\ 0, & \text{其他},\end{cases}$$

(1) 求常数 c;

(2) 求 $P\{X\leqslant 0.5,Y\leqslant 0.5\}$;

(3) 求 $P\{X+Y\leqslant 1\}$;

(4) 求 $P\{X>0.5\}$.

解: (1) 由联合密度函数性质

$$\begin{aligned}1&=\int_{-\infty}^{+\infty}\mathrm{d}x\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}y=\int_0^1\mathrm{d}x\int_0^1(c+xy)\mathrm{d}y\\&=\int_0^1\left(cy+\frac{1}{2}xy^2\right)\Big|_0^1\mathrm{d}x=\int_0^1\left(c+\frac{1}{2}x\right)\mathrm{d}x\\&=\left(cx+\frac{1}{4}x^2\right)\Big|_0^1=c+\frac{1}{4}.\end{aligned}$$

因此 $c=\dfrac{3}{4}$.

(2)

$$\begin{aligned}P\{X\leqslant 0.5,Y\leqslant 0.5\}&=\int_0^{0.5}\mathrm{d}x\int_0^{0.5}\left(\frac{3}{4}+xy\right)\mathrm{d}y=\int_0^{0.5}\left(\frac{3}{4}y+\frac{1}{2}xy^2\right)\Big|_0^{0.5}\mathrm{d}x\\&=\int_0^{0.5}\left(\frac{3}{8}+\frac{1}{8}x\right)\mathrm{d}x=\left(\frac{3}{8}x+\frac{1}{16}x^2\right)\Big|_0^{0.5}\\&=\frac{3}{16}+\frac{1}{64}=\frac{13}{64}.\end{aligned}$$

(3)

$$P\{X+Y\leqslant 1\}=\int_0^1\mathrm{d}x\int_0^{1-x}\left(\frac{3}{4}+xy\right)\mathrm{d}y=\int_0^1\left(\frac{3}{4}y+\frac{1}{2}xy^2\right)\Big|_0^{1-x}\mathrm{d}x$$

$$= \int_0^1 \left[\frac{3}{4}(1-x) + \frac{1}{2}x(1-x)^2\right] \mathrm{d}x = \int_0^1 \left(\frac{1}{2}x^3 - x^2 - \frac{1}{4}x + \frac{3}{4}\right) \mathrm{d}x$$
$$= \left(\frac{1}{8}x^4 - \frac{1}{3}x^3 - \frac{1}{8}x^2 + \frac{3}{4}x\right)\Big|_0^1 = \frac{1}{8} - \frac{1}{3} - \frac{1}{8} + \frac{3}{4} = \frac{5}{12}.$$

(4) 由于 $P\{X > 0.5\} = 1 - P\{X \leqslant 0.5\}$, 且

$$\begin{aligned} P\{X \leqslant 0.5\} &= \int_0^{0.5} \mathrm{d}x \int_0^1 \left(\frac{3}{4} + xy\right) \mathrm{d}y = \int_0^{0.5} \left(\frac{3}{4}y + \frac{1}{2}xy^2\right)\Big|_0^1 \mathrm{d}x \\ &= \int_0^{0.5} \left(\frac{3}{4} + \frac{1}{2}x\right) \mathrm{d}x = \left(\frac{3}{4}x + \frac{1}{4}x^2\right)\Big|_0^{0.5} \\ &= \frac{3}{8} + \frac{1}{16} = \frac{7}{16}. \end{aligned}$$

从而 $P\{X > 0.5\} = 1 - \frac{7}{16} = \frac{9}{16}$.

A15. 设二元连续型随机变量 (X, Y) 具有联合密度函数

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

(1) 分别求 X 及 Y 的边际密度函数;

(2) 求 $P\{Y \leqslant 2X\}$.

解: (1) X 的边际密度函数记为 $f_X(x)$.

当 $0 < x < 1$ 时, $f_X(x) = \int_{-\infty}^{+\infty} f(x,y)\mathrm{d}y = \int_0^2 x\mathrm{d}y = xy\Big|_0^2 = 2x$.

其他情况时, $f_X(x) = \int_0^2 0\,\mathrm{d}y = 0$.

从而

$$f_X(x) = \begin{cases} 2x, & 0 < x < 1, \\ 0, & \text{其他}. \end{cases}$$

Y 的边际密度函数记为 $f_Y(y)$.

当 $0 < y < 2$ 时, $f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)\mathrm{d}x = \int_0^1 x\mathrm{d}x = \frac{1}{2}x^2\Big|_0^1 = \frac{1}{2}$.

其他情况时, $f_Y(y) = \int_0^1 0\,\mathrm{d}x = 0$.

从而

$$f_Y(y) = \begin{cases} \frac{1}{2}, & 0 < y < 2, \\ 0, & \text{其他}. \end{cases}$$

(2)

$$\begin{aligned}P\{Y \leqslant 2X\} &= \int_0^1 \mathrm{d}x \int_0^{2x} x\mathrm{d}y = \int_0^1 xy\Big|_0^{2x} \mathrm{d}x \\ &= \int_0^1 2x^2 \mathrm{d}x = \frac{2}{3}x^3\Big|_0^1 = \frac{2}{3}.\end{aligned}$$

A16. 二元随机变量 (X, Y) 的联合密度函数为

$$f(x,y) = \begin{cases} c(x-1), & 1<x<2, x<y<4-x, \\ 0, & 其他. \end{cases}$$

(1) 求常数 c;

(2) 分别求 X, Y 的边际密度函数 $f_X(x)$ 与 $f_Y(y)$.

解: (1) 由题, $\int_{-\infty}^{+\infty} f(x,y)\mathrm{d}x\mathrm{d}y = 1$, 因此有

$$\begin{aligned}1 &= \int_1^2 \mathrm{d}x \int_x^{4-x} c(x-1)\mathrm{d}y = \int_1^2 c(x-1)[(4-x)-x]\mathrm{d}x = 2c\int_1^2 (-x^2+3x-2)\mathrm{d}x \\ &= 2c\left(-\frac{1}{3}x^3 + \frac{3}{2}x^2 - 2x\right)\Big|_1^2 = 2c\left[\left(-\frac{8}{3}+6-4\right) - \left(-\frac{1}{3}+\frac{3}{2}-2\right)\right] \\ &= 2c\left(-\frac{2}{3}+\frac{5}{6}\right) = \frac{c}{3}.\end{aligned}$$

因此 $c = 3$.

(2) 当 $1<x<2$ 时,

$$f_X(x) = \int_x^{4-x} 3(x-1)\mathrm{d}y = 3(x-1)[(4-x)-x] = -6(x-1)(x-2).$$

其他时 $f_X(x) = 0$, 则 $f_X(x) = \begin{cases} -6(x-1)(x-2), & 1<x<2, \\ 0, & 其他. \end{cases}$

由 $x<y<4-x$ 则 $x<4-y$. 当 $y=4-y$ 时 $y=2$. 又因为 $1<x<2$, 则 $1<y<3$.

当 $1<y<2$ 时, $x<y<4-y$, 此时

$$f_Y(y) = \int_1^y 3(x-1)\mathrm{d}x = \frac{3}{2}(x-1)^2\Big|_1^y = \frac{3}{2}(y-1)^2.$$

当 $2 \leqslant y < 3$ 时, $x < 4-y \leqslant y$, 此时

$$f_Y(y)=\int_1^{4-y}3(x-1)\mathrm{d}x=\frac{3}{2}(x-1)^2\Big|_1^{4-y}=\frac{3}{2}(4-y-1)^2=\frac{3}{2}(y-3)^2.$$

其他时 $f_Y(y)=0$, 则 $f_Y(y)=\begin{cases}\dfrac{3}{2}(y-1)^2, & 1<y<2,\\ \dfrac{3}{2}(y-3)^2 & 2\leqslant y<3,\\ 0, & 其他.\end{cases}$

A17. 二元随机变量 (X,Y) 的联合密度函数为

$$f(x,y)=\begin{cases}e^{-x}, & 0<y<x,\\ 0, & 其他.\end{cases}$$

(1) 分别求 X,Y 的边际密度函数 $f_X(x)$ 与 $f_Y(y)$;

(2) 求条件密度函数 $f_{Y|X}(y|x)$;

(3) 给定 $\{X=x\}$ 的条件下, Y 的条件分布是均匀分布吗? 为什么?

解: (1) $x>0$ 时, $f_X(x)=\int_0^x e^{-x}\mathrm{d}y=xe^{-x}$. 且 $x\leqslant 0$ 时 $f_X(x)=0$.

$y>0$ 时, $f_Y(y)=\int_y^{+\infty}e^{-x}\mathrm{d}x=-e^{-x}\Big|_y^{+\infty}=e^{-y}$. 且 $y\leqslant 0$ 时 $f_Y(y)=0$.

因此 $f_X(x)=\begin{cases}xe^{-x}, & 0<x<1,\\ 0, & 其他.\end{cases}$ $\quad f_Y(y)=\begin{cases}e^{-y}, & 0<y<1,\\ 0, & 其他.\end{cases}$

(2) 因为当 $0<y<x$ 时有 $f_{Y|X}(y|x)=\dfrac{f(x,y)}{f_X(x)}=\dfrac{1}{x}$, 其他情况时 $f_{Y|X}(y|x)=0$.

则有

$$f_{Y|X}(y|x)=\begin{cases}\dfrac{1}{x}, & 0<y<x,\\ 0, & 其他.\end{cases}$$

(3) 由 (2) 中条件密度函数可知, 当 $\{X=x\}$ 给定时, 把 x 看作常数, Y 的条件分布满足均匀分布.

A18. 设二维连续型随机变量 (X,Y) 的联合密度函数为

$$f(x,y)=\begin{cases}\dfrac{2x+y}{4}, & 0<x<1,0<y<2,\\ 0, & 其他,\end{cases}$$

(1) 分别求 X 与 Y 的边际密度函数;

(2) 分别求条件密度函数 $f_{X|Y}(x|y)$ 与 $f_{Y|X}(y|x)$;

(3) 分别求 $P\{X\leqslant 0.5|Y=0.5\}$ 与 $P\{Y\leqslant 0.5|X=0.5\}$.

解: (1) $0<x<1$ 时, $f_X(x)=\int_0^2\frac{2x+y}{4}\mathrm{d}y=\frac{y^2+4xy}{8}\Big|_0^2=\frac{4+8x}{8}=x+\frac{1}{2}$.

又其他情况下 $f_X(x)=0$, 从而有 $f_X(x)=\begin{cases} x+\frac{1}{2}, & 0<x<1, \\ 0, & \text{其他}. \end{cases}$

$0<y<2$ 时, $f_Y(y)=\int_0^1\frac{2x+y}{4}\mathrm{d}x=\frac{x^2+xy}{4}\Big|_0^1=\frac{y+1}{4}$. 又其他情况下 $f_Y(y)=0$.

从而有 $f_Y(y)=\begin{cases} \frac{y+1}{4}, & 0<y<2, \\ 0, & \text{其他}. \end{cases}$

(2) 当 $0<x<1$, $0<y<2$ 时,

$f_{X|Y}(x|y)=\frac{f(x,y)}{f_Y(y)}=\frac{\frac{2x+y}{4}}{\frac{y+1}{4}}=\frac{2x+y}{y+1}$, $f_{Y|X}(y|x)=\frac{f(x,y)}{f_X(x)}=\frac{\frac{2x+y}{4}}{x+\frac{1}{2}}=\frac{2x+y}{4x+2}$.

而其他情况时, $f_{X|Y}(x|y)=f_{Y|X}(y|x)=0$.

因此 $f_{X|Y}(x|y)=\begin{cases} \frac{2x+y}{y+1}, & 0<x<1, 0<y<2, \\ 0, & \text{其他}. \end{cases}$ $f_{Y|X}(y|x)=\begin{cases} \frac{2x+y}{4x+2}, & 0<x<1, 0<y<2, \\ 0, & \text{其他}. \end{cases}$

(3)

$$\begin{aligned} P\{X\leqslant 0.5|Y=0.5\} &= \int_0^{0.5} f_{X|Y}(x|0.5)\mathrm{d}x=\int_0^{0.5}\frac{2x+0.5}{1.5}\mathrm{d}x=\frac{x^2+0.5x}{1.5}\Big|_0^{0.5} \\ &= \frac{0.5^2+0.5\times 0.5}{1.5}=\frac{0.5}{1.5}=\frac{1}{3}. \end{aligned}$$

$$\begin{aligned} P\{Y\leqslant 0.5|X=0.5\} &= \int_0^{0.5} f_{Y|X}(y|0.5)\mathrm{d}y=\int_0^{0.5}\frac{1+y}{2+2}\mathrm{d}y=\frac{y^2+2y}{8}\Big|_0^{0.5} \\ &= \frac{0.5^2+2\times 0.5}{8}=\frac{1.25}{8}=\frac{5}{32}. \end{aligned}$$

A19. 设 (X,Y) 为二元随机变量, X 的密度函数为

$$f_X(x)=\begin{cases} \lambda^2 x e^{-\lambda x}, & x>0, \\ 0, & x\leqslant 0. \end{cases}$$

其中 $\lambda>0$. 当 $x>0$ 时, 给定 $\{X=x\}$ 的条件下 Y 的条件密度函数为

$$f_{Y|X}(y|x)=\begin{cases} \frac{1}{x}e^{-\frac{y}{x}}, & y>0, \\ 0, & y\leqslant 0. \end{cases}$$

(1) 求 (X,Y) 的联合密度函数;

(2) 求当 $x>0$ 时, 给定 $\{X=x\}$ 的条件下 Y 的条件分布函数;

(3) 求 $P\{Y>1|X=1\}$.

解: (1) 因为联合密度函数 $f(x,y)=f_X(x)\cdot f_{Y|X}(y|x)$, 从而有

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

(2) 此时条件分布函数 $F_{Y|X}(y|x)=P\{Y\leqslant y|X=x\}$, 从而 $y\leqslant 0$ 时 $F_{Y|X}(y|x)=0$.

$y>0$ 时,

$$F_{Y|X}(y|x)=\int_0^y f_{Y|X}(t|x)\mathrm{d}t=\int_0^y \frac{1}{x}e^{-\frac{t}{x}}\mathrm{d}t=-\int_0^y \mathrm{d}e^{-\frac{t}{x}}=-e^{-\frac{t}{x}}\Big|_0^y=1-e^{-\frac{y}{x}}.$$

从而 $x>0$ 时有

$$F_{Y|X}(y|x)=\begin{cases}1-e^{-\frac{y}{x}}, & y>0,\\ 0, & y\leqslant 0.\end{cases}$$

(3) 由(2), $F_{Y|X}(1|1)=P\{Y\leqslant 1|X=1\}=1-e^{-1}$.

因此 $P\{Y>1|X=1\}=1-P\{Y\leqslant 1|X=1\}=1-(1-e^{-1})=e^{-1}$.

A20. 设二元随机变量 (X,Y) 的联合密度函数为

$$f(x,y)=\begin{cases}\frac{5}{4}x, & y^2<x<1,\\ 0, & \text{其他}.\end{cases}$$

(1) 求 Y 的边际密度函数 $f_Y(y)$;

(2) 求条件密度函数 $f_{X|Y}(x|y)$;

(3) 求 $P\left\{X>\frac{1}{2}\middle|Y=\frac{1}{2}\right\}$.

解: (1) 由题, 当 $y^2<1$ 时有 $-1<y<1$, 此时有 $f_Y(y)=\int_{y^2}^1\frac{5}{4}x\mathrm{d}x=\frac{5}{8}x^2\Big|_{y^2}^1=\frac{5}{8}-\frac{5}{8}y^4$.

从而 $f_Y(y)=\begin{cases}\frac{5}{8}-\frac{5}{8}y^4, & -1<y<1,\\ 0, & \text{其他}.\end{cases}$

(2) $-1<y<1$ 且 $y^2<x<1$ 时, $f_{X|Y}(x|y)=\dfrac{f(x,y)}{f_Y(y)}=\dfrac{\frac{5}{4}x}{\frac{5}{8}-\frac{5}{8}y^4}=\dfrac{2x}{1-y^4}$.

因此 $f_{X|Y}(x|y)=\begin{cases}\dfrac{2x}{1-y^4}, & y^2<x<1,\\ 0, & 其他.\end{cases}$

(3)

$$P\left\{X>\frac{1}{2}\middle|Y=\frac{1}{2}\right\}=\int_{\frac{1}{2}}^{1}f_{X|Y}\left(x|\tfrac{1}{2}\right)\mathrm{d}x=\int_{\frac{1}{2}}^{1}\frac{2x}{1-\left(\frac{1}{2}\right)^4}\mathrm{d}x=\frac{16}{15}x^2\bigg|_{\frac{1}{2}}^{1}=\frac{16}{15}\left(1-\frac{1}{4}\right)=\frac{4}{5}.$$

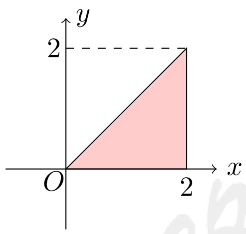
A21. 设随机变量 (X,Y) 服从以 $(0,0)$, $(2,0)$, $(2,2)$ 为顶点的三角形区域上均匀分布.

(1) 求 (X,Y) 的联合密度函数;

(2) 求 $P\{X+Y>2\}$;

(3) 求 $P\{X<1\}$.

解: 三角形区域内部 $D=\{(x,y)|0<x<2,0<y<x\}$. 如下图红色区域所示.

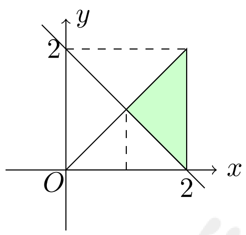


其面积为 $S=\dfrac{1}{2}\times2\times2=2$.

(1) 由均匀分布, 联合密度函数为

$$f(x,y)=\begin{cases}\dfrac{1}{S}=\dfrac{1}{2}, & 0<y<x<2,\\ 0, & 其他.\end{cases}$$

(2) 满足 $X+Y>2$ 的区域 $D_1=\{(x,y)|0<y<x<2,x+y>2\}$. 如下图绿色区域所示.



因为直线 $y=x$ 与 $x+y=2$ 交点为 $(1,1)$, 此时 D_1 的面积 $S_1=\dfrac{1}{2}\times1\times2=1$.

从而 $P\{X+Y>2\}=\dfrac{S_1}{S}=\dfrac{1}{2}$.

(3) 满足 $X<1$ 的区域 $D_2=\{0<x<1,0<y<x\}$. 此时有

$$P\{X<1\}=\int_0^1\mathrm{d}x\int_0^x f(x,y)\mathrm{d}y=\int_0^1\frac{1}{2}x\mathrm{d}x=\frac{1}{4}x^2\bigg|_0^1=\frac{1}{4}.$$

A22. 在区间 $(0,1)$ 内随机取一数 X, 在 $\{X=x\}$ 的条件下再在区间 $(x,1)$ 内随机取一数 Y.

(1) 求 (X,Y) 的联合密度函数;

(2) 求给定 $\{Y=y\}(0<y<1)$ 的条件下 X 的条件密度函数.

解: (1) 由题, X 的密度函数 $f_X(x)=\begin{cases}1, & 0<x<1,\\ 0, & 其他.\end{cases}$

且此时在 $\{X=x\}$ 的条件下, $f_{Y|X}(y|x)=\begin{cases}\dfrac{1}{1-x}, & 0<x<y<1,\\ 0, & 其他.\end{cases}$

因此联合密度函数 $f(x,y)=f_X(x)\cdot f_{Y|X}(y|x)=\begin{cases}\dfrac{1}{1-x}, & 0<x<y<1,\\ 0, & 其他.\end{cases}$

(2) 因为此时 $0<x<y$, 在 $0<y<1$ 时 $f_Y(y)=\displaystyle\int_0^y \frac{1}{1-x}\mathrm{d}x=-\ln(1-x)\Big|_0^y=-\ln(1-y)$.

则 $0<x<y<1$ 时 $f_{Y|X}(x|y)=\dfrac{f(x,y)}{f_Y(y)}=\dfrac{1}{-(1-x)\ln(1-y)}$.

从而条件密度函数 $f_{Y|X}(x|y)=\begin{cases}\dfrac{1}{-(1-x)\ln(1-y)}, & 0<x<y<1,\\ 0, & 其他.\end{cases}$

A23. 设二维随机变量 (X,Y) 服从分布 $N(\mu_1,\mu_2;\sigma_1^2,\sigma_2^2;\rho)$, 其中 $\mu_1=0$, $\mu_2=1$, $\sigma_1^2=1$, $\sigma_2^2=2$, $\rho=-\dfrac{1}{2}$.

(1) 写出 X,Y 各自的边际密度函数;

(2) 写出给定 $\{X=0\}$ 的条件下 Y 的条件密度函数;

(3) 求 $P\{Y\leqslant 1|X=0\}$.

解: (1) 根据二元正态分布可知, $X\sim N(0,1)$, $Y\sim N(1,2)$.

因此边际密度函数 $f_X(x)=\dfrac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}$, $x\in\mathbb{R}$; $\quad f_Y(y)=\dfrac{1}{\sqrt{4\pi}}\mathrm{e}^{-\frac{(y-1)^2}{4}}$, $y\in\mathbb{R}$.

(2) 根据二元正态分布可知, 当 $\{X=0\}$ 时, Y 的条件分布也为正态分布.

此时 $Y\sim N\left(\mu_2+\rho\dfrac{\sigma_2}{\sigma_1}(0-\mu_1),(1-\rho^2)\sigma_2^2\right)$, 即 $Y\sim N\left(1,\dfrac{3}{2}\right)$.

从而给定 $\{X=0\}$ 的条件下 Y 的条件密度函数 $f_{Y|X}(y|0)=\dfrac{1}{\sqrt{3\pi}}\mathrm{e}^{-\frac{(y-1)^2}{3}}$, $y\in\mathbb{R}$.

(3)

$$\begin{aligned}P\{Y\leqslant 1|X=0\}&=\int_{-\infty}^1 f_{Y|X}(y|0)\mathrm{d}y\xlongequal{u=y-1}\int_{-\infty}^0\frac{1}{\sqrt{3\pi}}\mathrm{e}^{-\frac{u^2}{3}}\mathrm{d}u\\&=\frac{1}{2}\int_{-\infty}^{+\infty}\frac{1}{\sqrt{3\pi}}\mathrm{e}^{-\frac{u^2}{3}}\mathrm{d}u=\frac{1}{2}\cdot 1=\frac{1}{2}.\end{aligned}$$

A24. 二维离散型随机变量 (X,Y) 的联合分布律为

X	Y		
	0	1	2
0	0.1	0.1	0.1
1	0.2	0.2	0.2
2	0.1	0	0

判断 X 与 Y 是否相互独立, 并说明理由.

解: 因为 $P\{X=2\}=0.1$, $P\{Y=2\}=0.1+0.2=0.3$,

但 $P\{X=2,Y=2\}=0\neq P\{X=2\}\cdot P\{Y=2\}$.

从而 X 与 Y 不相互独立.

A25. 设二维离散型随机变量 (X,Y) 的联合分布律为

X	Y		
	0	1	2
0	0.1	0.2	0.1
1	0.05	0.1	0.05
2	a	b	c

已知 X 与 Y 相互独立, 求 a,b,c.

解: 由题, $P\{Y=0\}=0.1+0.05+a=0.15+a$; $P\{Y=1\}=0.2+0.1+b=0.3+b$;

$P\{Y=2\}=0.1+0.05+c=0.15+c$. 且 $0.15+a+0.3+b+0.15+c=1$, 即 $a+b+c=0.4$.

且 $P\{X=0\}=0.4$, $P\{X=1\}=0.2$, $P\{X=2\}=a+b+c=0.4$.

因为 X 与 Y 相互独立, 则 $P\{X=2,Y=0\}=P\{X=2\}\cdot P\{Y=0\}=0.4(0.15+a)=a$, 解得 $a=0.1$.

$P\{X=2,Y=1\}=P\{X=2\}\cdot P\{Y=1\}=0.4(0.3+b)=b$, 解得 $b=0.2$.

所以 $c=0.4-a-b=0.1$.

A26. 二维连续型随机变量 (X,Y) 的联合密度函数为下列各式时, 判断对应的 X 与 Y 是否相互独立, 并说明理由:

$$(1)\ f(x,y)=\begin{cases}\dfrac{2x+y}{4}, & 0<x<1,0<y<2,\\ 0, & \text{其他};\end{cases}$$

$$(2)\ f(x,y)=\begin{cases} xy, & 0<x<1, 0<y<2, \\ 0, & \text{其他}; \end{cases}$$

$$(3)\ f(x,y)=\begin{cases} \dfrac{3x}{4}, & 0<x<y<2, \\ 0, & \text{其他}. \end{cases}$$

解: (1) 当 $0<x<1$ 时,

$$f_X(x)=\int_{-\infty}^{+\infty} f(x,y)\mathrm{d}y=\int_0^2 \frac{2x+y}{4}\mathrm{d}y=\left.\frac{4xy+y^2}{8}\right|_0^2=\frac{8x+4-0}{8}=x+\frac{1}{2}.$$

其他情况下 $f_X(x)=0$, 则 $f_X(x)=\begin{cases} x+\dfrac{1}{2}, & 0<x<1, \\ 0, & \text{其他}. \end{cases}$

当 $0<y<2$ 时,

$$f_Y(y)=\int_{-\infty}^{+\infty} f(x,y)\mathrm{d}x=\int_0^1 \frac{2x+y}{4}\mathrm{d}x=\left.\frac{x^2+xy}{4}\right|_0^1=\frac{1+y-0}{4}=\frac{y}{4}+\frac{1}{4}.$$

其他情况下 $f_Y(y)=0$, 则 $f_Y(y)=\begin{cases} \dfrac{y}{4}+\dfrac{1}{4}, & 0<y<2, \\ 0, & \text{其他}. \end{cases}$

从而 $0<x<1,\ 0<y<2$ 时, $f_X(x)\cdot f_Y(y)=\dfrac{xy}{4}+\dfrac{x}{4}+\dfrac{y}{8}+\dfrac{1}{8}$.

此时 $f(x,y)\neq f_X(x)\cdot f_Y(y)$, 从而 X 与 Y 不独立.

A27. 在半圆 $D=\{(x,y): x^2+y^2\leqslant 1, x>0\}$ 内随机投点 A, 设 A 点的坐标为 (X,Y).

(1) 求 X 的边际密度函数 $f_X(x)$;

(2) 求 $P\left\{X<\dfrac{1}{2}\right\}$;

(3) X 与 Y 相互独立吗？为什么？

解: (1) 由题 (X,Y) 服从半圆 D 上的均匀分布.

又因为半圆 D 的面积为 $\dfrac{\pi}{2}$, 故联合密度函数 $f(x,y)=\begin{cases} \dfrac{2}{\pi}, & x^2+y^2\leqslant 1, x>0, \\ 0, & \text{其他}. \end{cases}$

当 $x^2+y^2\leqslant 1,\ x>0$ 时, $0<x\leqslant 1$. 且此时有 $y^2\leqslant 1-x^2$, $-\sqrt{1-x^2}\leqslant y\leqslant \sqrt{1-x^2}$.

因此 $0<x<1$时, $f_X(x)=\displaystyle\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{2}{\pi}\mathrm{d}y=\frac{4}{\pi}\sqrt{1-x^2}$. 其他情况下 $f_X(x)=0$.

因此 $f_X(x)=\begin{cases}\dfrac{4}{\pi}\sqrt{1-x^2}, & 0<x<1,\\ 0, & \text{其他}.\end{cases}$

(2)

$$
\begin{aligned}
P\left\{X<\frac{1}{2}\right\}&=\int_0^{\frac{1}{2}}\frac{4}{\pi}\sqrt{1-x^2}\mathrm{d}x \xlongequal{x=\sin t}\frac{4}{\pi}\int_0^{\frac{\pi}{6}}\cos t\mathrm{d}\sin t=\frac{4}{\pi}\int_0^{\frac{\pi}{6}}\cos^2 t\mathrm{d}t\\
&=\frac{4}{\pi}\int_0^{\frac{\pi}{6}}\frac{1+\cos 2t}{2}\mathrm{d}t=\frac{4}{\pi}\left(\frac{t}{2}+\frac{\sin 2t}{4}\right)\Bigg|_0^{\frac{\pi}{6}}\\
&=\frac{4}{\pi}\left(\frac{\pi}{12}+\frac{1}{4}\sin\frac{\pi}{3}\right)=\frac{\sqrt{3}}{2\pi}+\frac{1}{3}.
\end{aligned}
$$

(3) 因为在该半圆内有 $-1\leqslant y\leqslant 1$, 且此时有 $0<x\leqslant\sqrt{1-y^2}$.

从而 $-1<y<1$ 时 $f_Y(y)=\displaystyle\int_0^{\sqrt{1-y^2}}\frac{2}{\pi}\mathrm{d}x=\frac{2}{\pi}\sqrt{1-y^2}$.

则 $x^2+y^2\leqslant 1$, $x>0$ 时, $f(x,y)\neq f_X(x)\cdot f_Y(y)$, 故 X,Y 不相互独立.

A28. 设二维随机变量 (X,Y) 服从分布 $N(0,1;2,4;0)$, 分别求 X 与 Y 的边际密度函数, 并判断 X 与 Y 是否相互独立.

解: 由题, $X\sim N(0,2)$, $Y\sim N(1,4)$.

从而 $f_X(x)=\dfrac{1}{\sqrt{2\pi}\cdot\sqrt{2}}e^{-\frac{x^2}{2\times 2}}=\dfrac{1}{2\sqrt{\pi}}e^{-\frac{x^2}{4}}$, $f_Y(y)=\dfrac{1}{\sqrt{2\pi}\cdot 2}e^{-\frac{(y-1)^2}{2\times 4}}=\dfrac{1}{2\sqrt{2\pi}}e^{-\frac{(y-1)^2}{8}}$.

因为 $\rho=0$, 从而 X 与 Y 相互独立.

A29. 设二维随机变量 (X,Y) 的联合密度函数为

$$
f(x,y)=\frac{1}{2}[f_1(x,y)+f_2(x,y)],
$$

其中 $f_1(x,y)$ 与 $f_2(x,y)$ 为两个二维正态变量 (X_1,Y_1) 与 (X_2,Y_2) 的联合密度函数, 且已知 (X_i,Y_i) $(i=1,2)$ 的边际分布均为标准正态分布.

(1) 求 X 与 Y 的边际密度函数 $f_X(x)$ 与 $f_Y(y)$;

(2) 当 (X_i,Y_i) 的分布中的参数 $\rho_i=0(i=1,2)$ 时, X 与 Y 相互独立吗?

解: 由题 $f_{X_1}(x)=f_{X_2}(x)=\dfrac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}$, $f_{Y_1}(y)=f_{Y_2}(y)=\dfrac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{y^2}{2}}$.

(1)

$$
\begin{aligned}
f_X(x)&=\int_{-\infty}^{+\infty}\frac{1}{2}[f_1(x,y)+f_2(x,y)]\mathrm{d}y=\frac{1}{2}\left[\int_{-\infty}^{+\infty}f_1(x,y)\mathrm{d}y+\int_{-\infty}^{+\infty}f_2(x,y)\mathrm{d}y\right]\\
&=\frac{1}{2}[f_{X_1}(x)+f_{X_2}(x)]=\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}},\qquad x\in\mathbb{R}.
\end{aligned}
$$

类似地, $f_Y(y)=\dfrac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{y^2}{2}}$, $y\in\mathbb{R}$.

(2) 因为 $\rho_1=0$, 则 $f_1(x,y)=f_{X_1}(x)\cdot f_{Y_1}(y)=\dfrac{1}{2\pi}\mathrm{e}^{-\frac{x^2+y^2}{2}}$.

又因为 $\rho_2=0$, 则 $f_2(x,y)=f_{X_2}(x)\cdot f_{Y_2}(y)=\dfrac{1}{2\pi}\mathrm{e}^{-\frac{x^2+y^2}{2}}$.

因此 $f(x,y)=\dfrac{1}{2}[f_1(x,y)+f_2(x,y)]=\dfrac{1}{2\pi}\mathrm{e}^{-\frac{x^2+y^2}{2}}=f_X(x)\cdot f_Y(y)$.

故 X,Y 相互独立.

A30. 设随机变量 X 与 Y 相互独立, $X\sim B(1,0.4)$, $Y\sim B(2,0.4)$, 令 $Z=X+Y$, 求 Z 的概率分布律.

解: 因为 X 可能取值为 $0,1$, Y 可能取值为 $0,1,2$.

从而 Z 可能取值为 $0,1,2,3$.

$Z=0$ 时, 则 $P\{Z=0\}=P\{X=0,Y=0\}=P\{X=0\}\cdot P\{Y=0\}=0.6\times0.6\times0.6=0.216$.

$Z=1$ 时,

$$\begin{aligned}P\{Z=1\}&=P\{X=0,Y=1\}+P\{X=1,Y=0\}\\&=P\{X=0\}\cdot P\{Y=1\}+P\{X=1\}\cdot P\{Y=0\}\\&=0.6\times\mathrm{C}_2^1\times0.4\times(1-0.4)+0.4\times0.6\times0.6=0.288+0.144=0.432.\end{aligned}$$

$Z=2$ 时,

$$\begin{aligned}P\{Z=2\}&=P\{X=0,Y=2\}+P\{X=1,Y=1\}\\&=P\{X=0\}\cdot P\{Y=2\}+P\{X=1\}\cdot P\{Y=1\}\\&=0.6\times0.4\times0.4+0.4\times\mathrm{C}_2^1\times0.6\times0.4=0.096+0.192=0.288.\end{aligned}$$

$Z=3$ 时, $P\{Z=3\}=P\{X=1,Y=2\}=P\{X=1\}\cdot P\{Y=2\}=0.4\times0.4\times0.4=0.064$.

A31. 某人连续参加 2 场比赛, 第 1,2 场比赛可得的奖金数分别为 X, Y, 且已知 X 的概率分布律为

X	0	1000	5000
p	0.5	0.3	0.2

Y 具有密度函数 $f(y)$, X 与 Y 相互独立, 求此人可得的奖金总数 $Z=X+Y$ 的密度函数.

解: 记 Y 的分布函数为 $F_Y(y)$, Z 的分布函数为 $F_Z(z)$. 由 X 与 Y 独立可知

$$\begin{aligned}F_Z(z)&=P\{Z\leqslant z\}=P\{X+Y\leqslant z\}\\&=P\{X=0,Y\leqslant z\}+P\{X=1000,Y\leqslant z-1000\}+P\{X=5000,Y\leqslant z-5000\}\end{aligned}$$

$$= P\{X=0\}P\{Y \leqslant z\} + P\{X=1000\}P\{Y \leqslant z-1000\} + P\{X=5000\}P\{Y \leqslant z-5000\}$$
$$= 0.5F_Y(z) + 0.3F_Y(z-1000) + 0.2F_Y(z-5000).$$

从而 Z 的密度函数 $f_Z(z) = F_Z'(z) = 0.5f(z) + 0.3f(z-1000) + 0.2f(z-5000)$.

A32. 设随机变量 X 与 Y 相互独立, 且分别具有概率分布律

X	0	1	2
p	0.2	0.3	0.5

Y	1	2	3
p	0.2	0.4	0.4

设 $Z = X+Y$, $M = \max\{X,Y\}$, $N = \min\{X,Y\}$, 分别求 Z, M, N 的概率分布律.

解: Z 可以取 $1,2,3,4,5$.

$$P\{Z=1\} = P\{X=0, Y=1\} = P\{X=0\} \cdot P\{Y=1\} = 0.04.$$

$$P\{Z=2\} = P\{X=0, Y=2\} + P\{X=1, Y=1\}$$
$$= P\{X=0\} \cdot P\{Y=2\} + P\{X=1\} \cdot P\{Y=1\}$$
$$= 0.08 + 0.06 = 0.14.$$

$$P\{Z=3\} = P\{X=0, Y=3\} + P\{X=1, Y=2\} + P\{X=2, Y=1\}$$
$$= P\{X=0\} \cdot P\{Y=3\} + P\{X=1\} \cdot P\{Y=2\} + P\{X=2\} \cdot P\{Y=1\}$$
$$= 0.08 + 0.12 + 0.1 = 0.3.$$

$$P\{Z=4\} = P\{X=1, Y=3\} + P\{X=2, Y=2\}$$
$$= P\{X=1\} \cdot P\{Y=3\} + P\{X=2\} \cdot P\{Y=2\}$$
$$= 0.12 + 0.2 = 0.32.$$

$$P\{Z=5\} = P\{X=2, Y=3\} = P\{X=2\} \cdot P\{Y=3\} = 0.2.$$

M 可以取 $1,2,3$.

$$P\{M=1\} = P\{X=0, Y=1\} + P\{X=1, Y=1\} = 0.04 + 0.06 = 0.1.$$

$$P\{M=2\} = P\{X=2, Y=1\} + P\{Y=2\} = 0.1 + 0.4 = 0.5.$$

$$P\{M=3\} = P\{Y=3\} = 0.4.$$

N 可以取 $0,1,2$.

$$P\{N=0\}=P\{X=0\}=0.2.$$

$$P\{N=1\}=P\{X=1\}+P\{X=2,Y=1\}=0.3+0.1=0.4.$$

$$P\{N=2\}=P\{X=2,Y=2\}+P\{X=2,Y=3\}=0.2+0.2=0.4.$$

A33. 设二维离散型随机变量 (X,Y) 的联合分布律为

X \ Y	0	1	2
1	0.2	0.1	0.3
2	0.2	0	0.2

设 $Z=XY$, $M=\max\{X,Y\}$, $N=\min\{X,Y\}$, 分别求 Z, M, N 的概率分布律.

解: Z 可以取 $0,1,2,4$.

$Z=0$ 时, $P\{Z=0\}=P\{Y=0\}=0.4$.

$Z=1$ 时, $P\{Z=1\}=P\{X=1,Y=1\}=0.1$.

$Z=2$ 时, $P\{Z=2\}=P\{X=1,Y=2\}+P\{X=2,Y=1\}=0.3$.

$Z=4$ 时, $P\{Z=4\}=P\{X=2,Y=2\}=0.2$.

M 可以取 $1,2$.

$M=1$ 时 $P\{M=1\}=P\{X=1,Y\leqslant 1\}=P\{X=1,Y=0\}+P\{X=1,Y=1\}=0.3$.

$M=2$ 时 $P\{M=2\}=1-P\{M=1\}=0.7$.

N 可以取 $0,1,2$.

$N=0$ 时 $P\{N=0\}=P\{Y=0\}=0.4$.

$N=1$ 时 $P\{N=1\}=P\{X=1,Y=1\}+P\{X=1,Y=2\}+P\{X=2,Y=1\}=0.4$.

$N=2$ 时 $P\{N=2\}=P\{X=2,Y=2\}=0.2$.

A34. 设随机变量 X 与 Y 相互独立且都服从参数为 1 的指数分布. 设 $Z=X+Y$, $M=\max\{X,Y\}$, $N=\min\{X,Y\}$, 分别求 Z, M, N 的密度函数.

解: 由题, X 的密度函数为 $f_X(x)=\begin{cases} e^{-x}, & x>0, \\ 0, & x\leqslant 0. \end{cases}$ Y 的密度函数为 $f_Y(y)=\begin{cases} e^{-y}, & y>0, \\ 0, & y\leqslant 0. \end{cases}$

$z \leqslant 0$ 时, $f_Z(z) = 0$. $z > 0$ 时,

$$\int_0^z f_X(x) f_Y(z-x) \mathrm{d}x = \int_0^z e^{-x} e^{-(z-x)} \mathrm{d}x = \int_0^z e^{-z} \mathrm{d}x = z e^{-z}.$$

因此 Z 的密度函数 $f_Z(z) = \begin{cases} z e^{-z}, & z > 0, \\ 0, & z \leqslant 0. \end{cases}$

因为 M 的分布函数 $F_M(t) = F_X(t) \cdot F_Y(t)$, 且 $t > 0$ 时, $F_X(t) = F_Y(t) = 1 - e^{-t}$.

因此 $t > 0$ 时 $F_M(t) = (1 - e^{-t})^2 = e^{-2t} - 2e^{-t} + 1$, 从而 $f_M(t) = F_M'(t) = -2e^{-2t} + 2e^{-t}$.

又 $t \leqslant 0$ 时 $f_M(t) = 0$, 则有 $f_M(t) = \begin{cases} -2e^{-2t} + 2e^{-t}, & t > 0, \\ 0, & t \leqslant 0. \end{cases}$

又因为 N 的分布函数 $F_N(t) = 1 - (1 - F_X(t))(1 - F_Y(t))$,

则 $t > 0$ 时 $F_N(t) = 1 - e^{-t} \cdot e^{-t} = 1 - e^{-2t}$; $t \leqslant 0$ 时 $F_N(t) = 0$.

因此有 $f_N(t) = F_N'(t) = \begin{cases} 2e^{-2t}, & t > 0, \\ 0, & t \leqslant 0. \end{cases}$

A35. 设随机变量 X 与 Y 相互独立且都服从标准正态分布.

(1) 设 $Z = X + Y$, 求 Z 的密度函数;

(2) 求 $P\{\min\{X, Y\} < 1\}$;

(3) 求 $P\{\max\{X, Y\} < 1\}$.

解: 由题, $X \sim N(0,1)$, $Y \sim N(0,1)$. 此时 X 与 Y 的密度函数 $f_X(t) = f_Y(t) = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$.

(1) 此时 $f_X(x) f_Y(z-x) = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \cdot \dfrac{1}{\sqrt{2\pi}} e^{-\frac{(z-x)^2}{2}} = \dfrac{1}{2\pi} e^{-\frac{2x^2 - 2zx + z^2}{2}} = \dfrac{1}{2\pi} e^{-\left(x - \frac{1}{2}z\right)^2} e^{-\frac{z^2}{4}}$.

所以有

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{+\infty} f_X(x) f_Y(z-x) \mathrm{d}x = \int_{-\infty}^{+\infty} \frac{1}{2\pi} e^{-\left(x - \frac{1}{2}z\right)^2} e^{-\frac{z^2}{4}} \mathrm{d}x \\ &\xlongequal{u = x - \frac{1}{2}z} \frac{1}{2\pi} e^{-\frac{z^2}{4}} \int_{-\infty}^{+\infty} e^{-u^2} \mathrm{d}u \xlongequal{t = \sqrt{2}u} \frac{1}{2\pi} e^{-\frac{z^2}{4}} \int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} \mathrm{d}\frac{t}{\sqrt{2}} \\ &= \frac{1}{2\sqrt{\pi}} e^{-\frac{z^2}{4}} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \mathrm{d}t = \frac{1}{2\sqrt{\pi}} e^{-\frac{z^2}{4}}. \end{aligned}$$

(2)

$$\begin{aligned} P\{\min\{X, Y\} < 1\} &= 1 - P\{\min\{X, Y\} \geqslant 1\} = 1 - P\{X \geqslant 1, Y \geqslant 1\} = 1 - P\{X \geqslant 1\} \cdot P\{Y \geqslant 1\} \\ &= 1 - (1 - P\{X < 1\})(1 - P\{Y < 1\}) = 1 - (1 - \Phi(1))^2 \approx 0.9748. \end{aligned}$$

(3)

$$P\{\max\{X,Y\}<1\}=P\{X<1,Y<1\}=P\{X<1\}\cdot P\{Y<1\}=(\Phi(1))^2\approx 0.7078.$$

A36. 设二维随机变量 (X,Y) 的联合密度函数为

$$f(x,y)=\begin{cases}\frac{1}{4}, & 0<x<2, 0<y<2x,\\ 0, & 其他.\end{cases}$$

记 $Z=2X-Y$, 求 Z 的密度函数.

解: 先求 Z 的分布函数.

因为 $0<x<2$, $0<y<2x$ 时有 $0<2x-y<4$, 由于 $F_Z(z)=P\{2X-Y\leqslant z\}$.

因此 $z\leqslant 0$ 时, $F_Z(z)=0$; 当 $z\geqslant 4$ 时, $F_Z(z)=1$.

当 $0<z<4$ 时, 【这里是二重积分的计算, 因为 $2x-z$ 与 0 大小未知所以是两个积分相加】

$$\begin{aligned}F_Z(z)=P\{2X-Y\leqslant z\}&=\iint\limits_{2x-y\leqslant z} f(x)\mathrm{d}x\mathrm{d}y=\int_{\frac{z}{2}}^{2}\mathrm{d}x\int_{2x-z}^{2x}\frac{1}{4}\mathrm{d}y+\int_0^{\frac{z}{2}}\mathrm{d}x\int_0^{2x}\frac{1}{4}\mathrm{d}y\\ &=\frac{1}{4}\int_{\frac{z}{2}}^{2}[2x-(2x-z)]\mathrm{d}x+\frac{1}{4}\int_0^{\frac{z}{2}}2x\mathrm{d}x\\ &=\frac{1}{4}z\left(2-\frac{z}{2}\right)+\frac{1}{4}x^2\Big|_0^{\frac{z}{2}}=\frac{z}{2}-\frac{z^2}{8}+\frac{1}{4}\cdot\frac{z^2}{4}=\frac{z}{2}-\frac{z^2}{16}.\end{aligned}$$

此时密度函数 $f_Z(z)=F_Z'(z)=\dfrac{1}{2}-\dfrac{z}{8}$. 且其他情况下密度函数为 0.

因此 Z 的密度函数 $f_Z(z)=\begin{cases}\frac{1}{2}-\frac{z}{8}, & 0<z<4,\\ 0, & 其他.\end{cases}$

A37. 设随机变量 X 与 Y 相互独立且都服从分布 $B(1,p)$ $(0<p<1)$, 定义

$$Z=\begin{cases}1, & X+Y=1,\\ 0, & X+Y\neq 1.\end{cases}$$

求 (X,Z) 的联合分布律.

解: 由题, $P\{X=1\}=P\{Y=1\}=p$, $P\{X=0\}=P\{Y=0\}=1-p$.

由独立知 $P\{X=0,Z=0\}=P\{X=0,Y=0\}=P\{X=0\}P\{Y=0\}=(1-p)^2$,

$P\{X=0,Z=1\}=P\{X=0,Y=1\}=P\{X=0\}P\{Y=1\}=p(1-p)$,

$P\{X=1,Z=1\}=P\{X=1,Y=0\}=P\{X=1\}P\{Y=0\}=p(1-p)$,

$P\{X=1,Z=0\}=P\{X=1,Y=1\}=P\{X=1\}P\{Y=1\}=p^2$.

从而 (X,Z) 的联合分布律为

X \ Z	0	1
0	$(1-p)^2$	$p(1-p)$
1	$p(1-p)$	p^2

A38. 设随机变量 $X\sim U(0,1)$, Y 的密度函数为

$$f_Y(y)=\begin{cases}2y, & 0<y<1,\\ 0, & \text{其他}.\end{cases}$$

且 X 与 Y 相互独立. 记 $M=\max\{X,Y\}$, $N=\min\{X,Y\}$, 分别求 M, N 的密度函数.

解: 由题, X 的密度函数为 $f_X(x)=\begin{cases}1, & 0<x<1,\\ 0, & \text{其他}.\end{cases}$

又因为 X 与 Y 相互独立, 从而 M 的分布函数 $F_M(t)=F_X(t)F_Y(t)=P\{X\leqslant t\}P\{Y\leqslant t\}$.

因为 $t\geqslant 1$ 时 $P\{X\leqslant t\}=P\{Y\leqslant t\}=1$;

$t\leqslant 0$ 时 $P\{X\leqslant t\}=P\{Y\leqslant t\}=0$;

当 $0<t<1$ 时, $P\{X\leqslant t\}=\int_0^t f_X(x)\mathrm{d}x=t$; $P\{Y\leqslant t\}=\int_0^t f_Y(y)\mathrm{d}y=y^2\Big|_0^t=t^2$.

因此

$$F_M(t)=\begin{cases}1, & t\geqslant 1,\\ t^3, & 0<t<1,\\ 0, & t\leqslant 0.\end{cases}\qquad \text{则 } f_M(t)=F_M'(t)=\begin{cases}3t^2, & 0<t<1,\\ 0, & \text{其他}.\end{cases}$$

又因为 N 的分布函数 $F_N(t)=1-(1-F_X(t))(1-F_Y(t))=1-(1-P\{X\leqslant t\})(1-P\{Y\leqslant t\})$.

由上可知, $t\geqslant 1$ 时, $F_N(t)=1-(1-1)(1-1)=1$;

$t\leqslant 0$ 时, $F_N(t)=1-(1-0)(1-0)=0$;

$0<t<1$ 时, $F_N(t)=1-(1-t)(1-t^2)=-t^3+t^2+t$.

则

$$f_N(t) = F_N'(t) = \begin{cases} -3t^2 + 2t + 1, & 0 < t < 1, \\ 0, & \text{其他}. \end{cases}$$

B1. 某公司为职工订报, 每位职工可以从 A, B, C 三种报中任订一份, 已知有 $\frac{2}{3}$ 的女职工决定订 A 报, 有 $\frac{3}{5}$ 的男职工决定订 B 报, 余下的人在三种报中随机选一份. 公司男、女职工各占一半, 从该公司中随机找一职工, 记

$$X=\begin{cases}1, & \text{此人为女职工},\\ 0, & \text{此人为男职工},\end{cases}\qquad Y=\begin{cases}1, & \text{此人订 A 报},\\ 2, & \text{此人订 B 报},\\ 3, & \text{此人订 C 报}.\end{cases}$$

(1) 写出 (X,Y) 的联合分布律;

(2) 求 Y 的边际分布律;

(3) 求给定 $\{Y=1\}$ 的条件下 X 的条件分布律.

解: (1) 由于余下的人选报纸是随机选取, 此时有

$P\{X=0,Y=1\}=\frac{1}{2}\times\frac{2}{5}\times\frac{1}{3}=\frac{1}{15}$;

$P\{X=0,Y=2\}=\frac{1}{2}\times\left(\frac{3}{5}+\frac{2}{5}\times\frac{1}{3}\right)=\frac{11}{30}$;

$P\{X=0,Y=3\}=\frac{1}{2}\times\frac{2}{5}\times\frac{1}{3}=\frac{1}{15}$;

$P\{X=1,Y=1\}=\frac{1}{2}\times\left(\frac{2}{3}+\frac{1}{3}\times\frac{1}{3}\right)=\frac{7}{18}$;

$P\{X=1,Y=2\}=\frac{1}{2}\times\frac{1}{3}\times\frac{1}{3}=\frac{1}{18}$;

$P\{X=1,Y=3\}=\frac{1}{2}\times\frac{1}{3}\times\frac{1}{3}=\frac{1}{18}$.

【这里要记得加上余下的人的订报】

从而 (X,Y) 的联合分布律为

X	Y		
	1	2	3
0	$\frac{1}{15}$	$\frac{11}{30}$	$\frac{1}{15}$
1	$\frac{7}{18}$	$\frac{1}{18}$	$\frac{1}{18}$

(2) 由(1)可知, $P\{Y=1\}=\frac{1}{15}+\frac{7}{18}=\frac{6+35}{90}=\frac{41}{90}$;

$P\{Y=2\}=\frac{11}{30}+\frac{1}{18}=\frac{33+5}{90}=\frac{38}{90}=\frac{19}{45}$;

$P\{Y=3\}=\frac{1}{15}+\frac{1}{18}=\frac{6+5}{90}=\frac{11}{90}$.

(3)

$$P\{X=0|Y=1\}=\frac{P\{X=0,Y=1\}}{P\{Y=1\}}=\frac{\frac{1}{15}}{\frac{41}{90}}=\frac{6}{41}.$$

$$P\{X=1|Y=1\}=\frac{P\{X=1,Y=1\}}{P\{Y=1\}}=\frac{\frac{7}{18}}{\frac{41}{90}}=\frac{35}{41}.$$

B2. 设某路段单位时间内发生的交通事故数 X 服从参数为 λ 的泊松分布, 其中事故原因是超速的概率为 0.1. 记因超速引发的事故数为 Y.

(1) 求 (X,Y) 的联合分布律;

(2) 求 Y 的边际分布律.

解: (1) 由题 $X\sim P(\lambda)$, 此时有 $P\{X=k\}=\frac{\mathrm{e}^{-\lambda}\lambda^k}{k!}$, $k=0,1,2,\cdots$.

又因为 $X=x$ 时, $Y\sim B(x,0.1)$. 从而

$$P\{X=i,Y=j\}=P\{X=i\}P\{Y=j|X=i\}=\frac{\mathrm{e}^{-\lambda}\lambda^i}{i!}\mathrm{C}_i^j0.1^j0.9^{i-j},$$

其中 $i=0,1,2,\cdots$, $j=0,1,\cdots i$.

(2) 对 $j=0,1,2,\cdots$,

$$\begin{aligned}P\{Y=j\}&=\sum_{i=0}^{+\infty}P\{X=i,Y=j\}=\sum_{i=0}^{+\infty}\frac{\mathrm{e}^{-\lambda}\lambda^i}{i!}\mathrm{C}_i^j0.1^j0.9^{i-j}=\sum_{i=j}^{+\infty}\frac{\mathrm{e}^{-\lambda}\lambda^i}{i!}\frac{i!}{j!(i-j)!}0.1^j0.9^{i-j}\\&=\frac{e^{-\lambda}\lambda^j0.1^j}{j!}\sum_{i=j}^{+\infty}\frac{(0.9\lambda)^{i-j}}{(i-j)!}=\frac{e^{-\lambda}(0.1\lambda)^j}{j!}\sum_{k=0}^{+\infty}\frac{(0.9\lambda)^k}{k!}\\&=\frac{e^{-\lambda}(0.1\lambda)^j}{j!}\cdot e^{0.9\lambda}=\frac{e^{-0.1\lambda}(0.1\lambda)^j}{j!}.\end{aligned}$$

其中最后一行利用了 e^x 的泰勒展开, 这里 $j=0,1,2,\cdots$.

B3. 设 (X,Y) 为二元随机变量, 已知 $P\{X=0,Y=0\}=P\{X=1,Y=1\}=0.1$, 现知 (X,Y) 落在 $D=\{(x,y):0<x<1,0<y<1\}$ 的任一小区域内的概率与该小区域面积成正比, 且 (X,Y) 只能落在点 $(0,0),(1,1)$ 及 D 内, 求 (X,Y) 的联合分布函数.

解: 由题, (X,Y) 落在 D 内的概率为 $1-0.1-0.1=0.8$.

又因为 D 的面积为 1, 由正比可得, 落在任一面积为 A 的区域的概率为 $0.8A$.

对联合分布函数 $F(x,y)=P\{X\leqslant x,Y\leqslant Y\}$, 当 $x<0$, $y<0$ 时, $F(x,y)=0$;

当 $x=0$, $y\geqslant 0$ 时, $F(x,y)=P\{X=0,Y=0\}=0.1$.

当 $y=0$, $x>0$ 时, $F(x,y)=P\{X=0,Y=0\}=0.1$.

【上述两种情况下面合并为 $x\geqslant 0$, $y\geqslant 0$ 且 $xy=0$.】

当 $0<x<1$, $0<y<1$ 时, 此时 (X,Y) 所在的小区域 $D_1=\{(m,n)|0<m\leqslant x,0<n\leqslant y\}$.

此时 D_1 面积为 xy, 从而落在 D_1 内的概率为 $0.8xy$. 此时有

$$F(x,y)=P\{X=0,Y=0\}+P\{0<X\leqslant x,0<Y\leqslant y\}=0.1+0.8xy.$$

当 $x\geqslant 1$, $0<y<1$ 时, $X\leqslant x$ 恒成立, 只需考虑 $Y\leqslant y$.

此时 $Y\leqslant y$ 的区域为 $D_2=\{(m,n)|0<m<1,0<n\leqslant y\}$, 其面积为 y.

从而

$$F(x,y)=P\{X=0,Y=0\}+P\{0<Y\leqslant y\}=0.1+0.8y.$$

当 $y\geqslant 1$, $0<x<1$ 时, $Y\leqslant y$ 恒成立, 只需考虑 $X\leqslant x$.

此时 $X\leqslant x$ 的区域为 $D_3=\{(m,n)|0<m\leqslant x,0<n<1\}$, 其面积为 x.

从而

$$F(x,y)=P\{X=0,Y=0\}+P\{0<X\leqslant x\}=0.1+0.8x.$$

当 $x\geqslant 1$, $y\geqslant 1$ 时, $X\leqslant x$ 与 $Y\leqslant y$ 恒成立, 有 $F(x,y)=1$.

综上则联合分布函数 $F(x,y)=\begin{cases}0, & x<0,y<0,\\ 0.1, & x\geqslant 0,y\geqslant 0 \text{ 且 } xy=0,\\ 0.1+0.8xy, & 0<x<1,0<y<1,\\ 0.1+0.8x, & 0<x<1,y\geqslant 1,\\ 0.1+0.8y, & x\geqslant 1,0<y<1,\\ 1, & x\geqslant 1,y\geqslant 1.\end{cases}$

B4. 设二元随机变量 (X,Y) 的联合密度函数为

$$f(x,y)=\begin{cases}c(y-x), & 0<x<y<1,\\ 0, & \text{其他}.\end{cases}$$

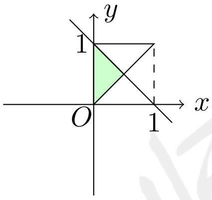
(1) 求常数 c;　　(2) 求 $P\{X+Y\leqslant 1\}$;　　(3) 求 $P\{X<0.5\}$.

解: (1) 由联合密度函数可知

$$\begin{aligned}1&=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x\mathrm{d}y=\int_0^1\mathrm{d}x\int_x^1c(y-x)\mathrm{d}y=c\int_0^1\left(\frac{1}{2}y^2-xy\right)\Big|_x^1\mathrm{d}x\\&=c\int_0^1\left[\left(\frac{1}{2}-x\right)-\left(\frac{1}{2}x^2-x^2\right)\right]\mathrm{d}x=c\int_0^1\left(\frac{1}{2}x^2-x+\frac{1}{2}\right)\mathrm{d}x\\&=c\left(\frac{1}{6}x^3-\frac{1}{2}x^2+\frac{1}{2}\right)\Big|_0^1=c\cdot\left(\frac{1}{6}-\frac{1}{2}+\frac{1}{2}\right)=\frac{c}{6}.\end{aligned}$$

所以 $c=6$.

(2)



由上图, 在三角形区域 $D:\{(x,y)|0<x<y<1\}$ 内联合密度函数不为 0.

且绿色区域 D_1 表示事件 $\{X+Y\leqslant 1\}$. 记 D_1 面积 S_1, D 的面积为 S.

由于 $S=\frac{1}{2}\times 1\times 1=\frac{1}{2}$, 且 $x+y=1$ 与 $y=x$ 交于点 $(0.5,0.5)$, 从而 $S_1=\frac{1}{2}\times 1\times 0.5=\frac{1}{4}$.

则 $P\{X+Y\leqslant 1\}=\frac{S_1}{S}=\frac{0.25}{0.5}=\frac{1}{2}$.

(3) 由题可得

$$\begin{aligned}
P\{X<0.5\}&=\int_0^{0.5}\mathrm{d}x\int_x^1 6(y-x)\mathrm{d}y=\int_0^{0.5}(3y^2-6xy)\Big|_x^1\mathrm{d}x\\
&=\int_0^{0.5}[(3-6x)-(3x^2-6x^2)]\mathrm{d}x=\int_0^{0.5}(3x^2-6x+3)\mathrm{d}x\\
&=(x^3-3x^2+3x)\Big|_0^{0.5}=\frac{1}{8}-\frac{3}{4}+\frac{3}{2}=\frac{7}{8}.
\end{aligned}$$

B5. 有一件工作需要甲、乙两人接力完成, 完成时间不超过 4 小时. 设甲先干了 X 小时, 再由乙完成, 加起来共用 Y 小时. 若 $X\sim U(1,2)$, 在 $\{X=x\}$ 条件下, Y 的条件密度函数

$$f_{Y|X}(y|x)=\begin{cases}\frac{2(4-y)}{(3-x)^2}, & x+1<y<4,\\ 0, & \text{其他}.\end{cases}$$

(1) 求 (X,Y) 的联合密度函数 $f(x,y)$ 及 $P\{Y<3\}$;

(2) 求 Y 的边际密度函数;

(3) 已知两人完成工作共花了 3 小时, 求甲的工作时间不超过 1.5 小时的概率.

解: (1) 由均匀分布, 则 X 的密度函数 $f_X(x)=\begin{cases}1, & 1<x<2,\\ 0, & \text{其他}.\end{cases}$

从而有 $f(x,y)=f_X(x)\cdot f_{Y|X}(y|x)=\begin{cases}\frac{2(4-y)}{(3-x)^2}, & 1<x<2, x+1<y<4,\\ 0, & \text{其他}.\end{cases}$

$$\begin{aligned}
P\{Y<3\}&=\int_1^2\mathrm{d}x\int_{x+1}^3\frac{2(4-y)}{(3-x)^2}\mathrm{d}y=\int_1^2\frac{8y-y^2}{(3-x)^2}\Big|_{x+1}^3\mathrm{d}x\\
&=\int_1^2\frac{(24-9)-[8(x+1)-(x+1)^2]}{(x-3)^2}\mathrm{d}x=\int_1^2\frac{x^2-6x+8}{(x-3)^2}\mathrm{d}x
\end{aligned}$$

$$= \int_1^2 \frac{(x-3)^2-1}{(x-3)^2}\mathrm{d}x = \int_1^2 \left[1-\frac{1}{(x-3)^2}\right]\mathrm{d}x = x+\frac{1}{x-3}\bigg|_1^2$$
$$= (2-1)-\left(1-\frac{1}{2}\right) = \frac{1}{2}.$$

(2) 由密度函数, $x<2$ 且 $x<y-1$. 此时应比较 $y-1$ 与 2 的大小关系.

$y-1\leqslant 2$ 时, $y\leqslant 3$, 又因为 $1<x<2$ 时有 $2<x+1<y$, 即此时 $2<y\leqslant 3$.

因为此时 $1<x<y-1$, 从而

$$f_Y(y) = \int_1^{y-1}\frac{2(4-y)}{(x-3)^2}\mathrm{d}x = \frac{2(y-4)}{x-3}\bigg|_1^{y-1} = \frac{2(y-4)}{y-1-3}-\frac{2(y-4)}{1-3} = 2+(y-4) = y-2.$$

$y-1>2$ 时, 有 $3<y<4$, 此时 $1<x<2$, 从而

$$f_Y(y) = \int_1^2 \frac{2(4-y)}{(x-3)^2}\mathrm{d}x = \frac{2(y-4)}{x-3}\bigg|_1^2 = \frac{2(y-4)}{2-3}-\frac{2(y-4)}{1-3} = 2(4-y)+y-4 = 4-y.$$

又因为其他情况密度函数为 0, 则有 $f_Y(y) = \begin{cases} y-2, & 2<y\leqslant 3, \\ 4-y, & 3<y<4, \\ 0, & \text{其他}. \end{cases}$

(3) 由题, 所求即为 $P\{X\leqslant 1.5|Y=3\}$.

从而给定 $\{Y=3\}$ 条件下 X 的条件密度函数 $f_{X|Y}(x|3) = \dfrac{f(x,3)}{f_Y(3)} = \dfrac{\frac{2(4-3)}{(3-x)^2}}{3-2} = \dfrac{2}{(x-3)^2}$, $1<x<2$.

则有

$$P\{X\leqslant 1.5|Y=3\} = \int_1^{1.5} f_{X|Y}(x|3)\mathrm{d}x = \int_1^{1.5}\frac{2}{(x-3)^2}\mathrm{d}x = \frac{2}{3-x}\bigg|_1^{1.5}$$
$$= \frac{2}{3-1.5}-\frac{2}{3-1} = \frac{4}{3}-1 = \frac{1}{3}.$$

B6. 在 A 地至 B 地 (距离为 m 公里) 的公路上, 事故发生地在离 A 地 X 公里处, 事故处理车在离 A 地 Y 公里处, X 与 Y 均服从 $(0,m)$ 上均匀分布, 且设 X 与 Y 相互独立. 求事故车与处理车的距离 Z 的密度函数.

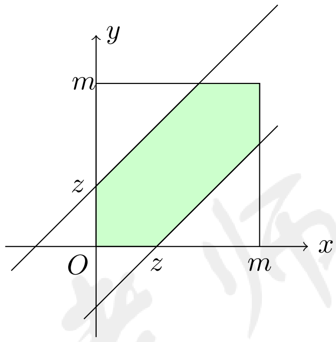
解: 由题, 此时事故车和处理车位于 A 地的同一端, 此时 $Z=|X-Y|$.

由于 X,Y 服从 $(0,m)$ 上均匀分布, 则 Z 只能在 0 到 m 上取值.

从而当 $z\leqslant 0$ 时, $F_Z(z)=P\{Z\leqslant z\}=0$; $z\geqslant m$ 时, $F_Z(z)=1$.

$0<z<m$ 时, 对 $|x-y|<z$ 有 $-z<y-x<z$, 即 $x-z<y<x+z$, 在下图中即绿色阴影部分.

记绿色部分为 D_1, 其面积为 S_1; 区域 $D=\{(x,y)|0<x<m, 0<y<m\}$, 其面积 S.



由均匀分布可知, $P\{Z \leqslant z\} = P\{|X-Y| \leqslant z\} = \dfrac{S_1}{S}$.

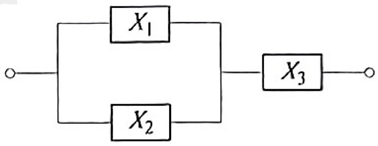
又因为 $S = m^2$, $S_1 = S - 2 \times \dfrac{1}{2}(m-z)(m-z) = m^2 - (m-z)^2 = 2mz - z^2$.

从而此时 $F_Z(z) = \dfrac{2mz - z^2}{m^2}$.

综上, Z 的分布函数 $F_Z(z) = \begin{cases} 0, & z \leqslant 0, \\ \dfrac{2mz-z^2}{m^2}, & 0 < z < m, \\ 1, & z \geqslant m. \end{cases}$

因此 Z 的密度函数 $f_Z(z) = F_Z'(z) = \begin{cases} \dfrac{2m-2z}{m^2}, & 0 < z < m, \\ 0, & \text{其他}. \end{cases}$

B7. 设一系统由三个相互独立的、正常工作时间分别为 X_1, X_2, X_3 的子系统组成(如图所示), 且 $X_i(i=1,2,3)$ 均服从参数为 λ 的指数分布, 求该系统正常工作时间 T 的分布函数 $F_T(t)$ 及密度函数 $f_T(t)$.



解: 系统正常工作时, 需要 X_3 正常工作且X_1, X_2 中至少有一个正常工作.

由于 $t>0$ 时 $P\{X_i \leqslant t\} = 1 - e^{-\lambda t}$, $i = 1,2,3$.

记 X_1, X_2 中至少有一个寿命大于 t 为事件 A.

则 $P(\overline{A}) = P\{X_1 \leqslant t, X_2 \leqslant t\} = (1-e^{-\lambda t})^2$, 因此 $P(A) = 1 - (1-e^{-\lambda t})^2$.

从而

$$P\{T > t\} = P\{X_3 > t\} \cdot P(A) = [1-(1-e^{-\lambda t})][1-(1-e^{-\lambda t})^2]$$

$$= e^{-\lambda t}(2e^{-\lambda t} - e^{-2\lambda t}) = 2e^{-2\lambda t} - e^{-3\lambda t}.$$

因此 $t>0$ 时分布函数 $F_T(t) = P\{T \leqslant t\} = 1 - P\{T > t\} = 1 - 2\mathrm{e}^{-2\lambda t} + \mathrm{e}^{-3\lambda t}$.

又因为 $t \leqslant 0$ 时 $F_T(t) = 0$, 则 $F_T(t) = \begin{cases} 1 - 2\mathrm{e}^{-2\lambda t} + \mathrm{e}^{-3\lambda t}, & t > 0, \\ 0, & t \leqslant 0. \end{cases}$

因此 $f_T(t)=F'_T(t)=\begin{cases}4\lambda\mathrm{e}^{-2\lambda t}-3\lambda\mathrm{e}^{-3\lambda t}, & t>0,\\ 0, & t\leqslant 0.\end{cases}$

B8. (1) 设随机变量 $X_1,X_2,\cdots,X_n$ 相互独立且均服从参数为 $p(0<p<1)$ 的 0-1 分布, 记 $Z=\sum\limits_{i=1}^{n}X_i$, 求 Z 的概率分布律;

(2) 设 $X\sim B(m,p),Y\sim B(n,p)$, 且 X 与 Y 相互独立. 记 $W=X+Y$, 求 W 的概率分布律.

解: (1) 对 $k=0,1,2,\cdots,n$, $\{Z=k\}$ 即表示 X_i $i=1,2,\cdots,n$ 中有 k 个数为 1, 另外 $n-k$ 个为 0.

因此 $P\{Z=k\}=\mathrm{C}_n^k p^k(1-p)^{n-k}$, 此时 $k=0,1,2,\cdots,n$.

(2) 因为 X 与 Y 相互独立, 且成功概率均为 p, 从而有 $W=X+Y\sim(m+n,p)$.

因此 $P\{W=k\}=\mathrm{C}_{m+n}^k p^k(1-p)^{m+n-k}$, 此时 $k=0,1,2,\cdots,m+n$.

B9. 设随机变量 X 服从区间 $(-a,a)$ 上均匀分布, 其中 $a>0$, $Y\sim N(\mu,\sigma^2)$, X 与 Y 相互独立, $Z=X+Y$, 求 Z 的密度函数.

解: 由题, X 的密度函数 $f_X(x)=\begin{cases}\dfrac{1}{2a}, & -a<x<a,\\ 0, & \text{其他}.\end{cases}$

又因为 $Y\sim N(\mu,\sigma^2)$, 则 Y 的密度函数 $f_Y(y)=\dfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$.

由于 X,Y 相互独立, 故

$$\begin{aligned}
f_Z(z)&=\int_{-\infty}^{+\infty}f_X(x)f_Y(z-x)\mathrm{d}x=\int_{-a}^{a}\frac{1}{2a}\cdot\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(z-x-\mu)^2}{2\sigma^2}}\mathrm{d}x\\
&=-\frac{1}{2a}\int_{-a}^{a}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\cdot\left(\frac{z-x-\mu}{\sigma}\right)^2}\mathrm{d}\frac{z-x-\mu}{\sigma}\xlongequal{u=\frac{z-x-\mu}{\sigma}}-\frac{1}{2a}\int_{\frac{z-a-\mu}{\sigma}}^{\frac{z+a-\mu}{\sigma}}\frac{1}{\sqrt{2\pi}}e^{-\frac{u^2}{2}}\mathrm{d}u\\
&=-\frac{1}{2a}\left[\Phi\left(\frac{z-a-\mu}{\sigma}\right)-\Phi\left(\frac{z+a-\mu}{\sigma}\right)\right]\\
&=\frac{1}{2a}\left[\Phi\left(\frac{z+a-\mu}{\sigma}\right)-\Phi\left(\frac{z-a-\mu}{\sigma}\right)\right].
\end{aligned}$$

其中 $\Phi(x)$ 为标准正态分布的分布函数.

B10. 设二元随机变量 (X,Y) 的联合密度函数为

$$f(x,y)=\begin{cases}\dfrac{3-x-y}{3}, & 0<x<1,0<y<2,\\ 0, & \text{其他}.\end{cases}$$

若 $Z=X+Y$, 求 Z 的密度函数.

解: 因为 $f_Z(z)=\int_{-\infty}^{+\infty} f(x,z-x)\mathrm{d}x$, 则要考虑 x 与 $z-x$ 的范围.

即 $0<x<1,\ 0<z-x<2$ 时, 有 $f(x,z-x)=\dfrac{3-x-(z-x)}{3}=\dfrac{3-z}{3}$.

又因为 $0<z-x<2$ 时, $z-2<x<z$, 从而要讨论 z 的取值.

$0<z\leqslant 1$ 时, $z-2<0<x<z\leqslant 1$, 有 $f_Z(z)=\int_0^z \dfrac{3-z}{3}\mathrm{d}x=z-\dfrac{z^2}{3}$.

$1<z\leqslant 2$ 时, $z-2\leqslant 0<x<1<z$, 有 $f_Z(z)=\int_0^1 \dfrac{3-z}{3}\mathrm{d}x=1-\dfrac{z}{3}$.

$2<z<3$ 时, $0<z-2<x<1<z$, 有 $f_Z(z)=\int_{z-2}^1 \dfrac{3-z}{3}\mathrm{d}x=\dfrac{(3-z)}{3}[1-(z-2)]=\dfrac{1}{3}z^2-2z+3$.

其他情况时, $f(x,z-x)=0$, 则 $f_Z(z)=0$.

综上: $f_Z(z)=\begin{cases} -\dfrac{z^2}{3}+z, & 0<z\leqslant 1, \\ -\dfrac{z}{3}+1, & 1<z\leqslant 2, \\ \dfrac{1}{3}z^2-2z+3, & 2<z<3, \\ 0, & \text{其他}. \end{cases}$

B11. 市场近期某种蔬菜的价格(单位: 元/公斤) $X\sim U(6,8)$(均匀分布), 某餐馆近期购买该种蔬菜的数量 Y 为 8 公斤和 10 公斤的概率均为 0.5. 求:

(1)购买金额 Z 不大于 60 元的概率 p;

(2)购买金额 Z 的分布函数 $F_Z(z)$.

解: 由题 X 的密度函数 $f_X(x)=\begin{cases} \dfrac{1}{2}, & 6<x<8, \\ 0, & \text{其他}. \end{cases}$ 且对 Y 有 $P\{Y=8\}=P\{Y=10\}=0.5$.

此时 X 与 Y 相互独立, 购买金额 $Z=XY$.

(1)由全概率公式可得

$$\begin{aligned} p=P\{Z\leqslant 60\}&=P\{XY\leqslant 60\}=P\{Y=8\}P\{XY\leqslant 60|Y=8\}+P\{Y=10\}P\{XY\leqslant 60|Y=10\} \\ &=0.5P\left\{X\leqslant \frac{15}{2}\right\}+0.5P\{X\leqslant 6\} \\ &=0.5\int_6^{\frac{15}{2}} \frac{1}{2}\mathrm{d}x+0=\frac{1}{2}\times\frac{1}{2}\times\left(\frac{15}{2}-6\right)=\frac{3}{8}. \end{aligned}$$

(2) 对 $6<X<8$, 由于 Y 取值只可能是 8 或 10, 从而 $48<XY<80$.

因此对 $F_Z(z)=P\{Z\leqslant z\}=P\{XY\leqslant z\}$, 当 $z\leqslant 48$ 时, $F_Z(z)=0$; 当 $z\geqslant 80$ 时, $F_Z(z)=1$;

又因为 $Y=8$ 时, $48<XY<64$; $y=10$ 时, $60<XY<80$.

从而当 $48<z<60$ 时, $\dfrac{z}{10}<6$, 有

$$F_Z(z)=P\{XY\leqslant z,Y=8\}+P\{XY\leqslant z,Y=10\}$$

$$=P\left\{X\leqslant\frac{z}{8}\right\}P\{Y=8\}+P\left\{X\leqslant\frac{z}{10}\right\}P\{Y=10\}$$
$$=\frac{1}{2}P\left\{X\leqslant\frac{z}{8}\right\}+0=\frac{1}{2}\int_6^{\frac{z}{8}}\frac{1}{2}\mathrm{d}x=\frac{1}{4}\left(\frac{z}{8}-6\right)=\frac{z}{32}-\frac{3}{2}.$$

当 $60\leqslant z<64$ 时,

$$\begin{aligned}F_Z(z)&=P\{XY\leqslant z,Y=8\}+P\{XY\leqslant z,Y=10\}\\&=P\left\{X\leqslant\frac{z}{8}\right\}P\{Y=8\}+P\left\{X\leqslant\frac{z}{10}\right\}P\{Y=10\}\\&=\frac{1}{2}P\left\{X\leqslant\frac{z}{8}\right\}+\frac{1}{2}P\left\{X\leqslant\frac{z}{10}\right\}\\&=\frac{1}{2}\int_6^{\frac{z}{8}}\frac{1}{2}\mathrm{d}x+\frac{1}{2}\int_6^{\frac{z}{10}}\frac{1}{2}\mathrm{d}x=\frac{1}{4}\left(\frac{z}{8}-6\right)+\frac{1}{4}\left(\frac{z}{10}-6\right)=\frac{9}{160}z-3.\end{aligned}$$

当 $64\leqslant z<80$ 时, $\frac{z}{8}\geqslant 8$, 有

$$\begin{aligned}F_Z(z)&=P\{XY\leqslant z,Y=8\}+P\{XY\leqslant z,Y=10\}\\&=P\left\{X\leqslant\frac{z}{8}\right\}P\{Y=8\}+P\left\{X\leqslant\frac{z}{10}\right\}P\{Y=10\}\\&=\frac{1}{2}+\frac{1}{2}P\left\{X\leqslant\frac{z}{8}\right\}=\frac{1}{2}+\frac{1}{2}\int_6^{\frac{z}{10}}\frac{1}{2}\mathrm{d}x=\frac{1}{2}+\frac{1}{4}\left(\frac{z}{10}-6\right)=\frac{z}{40}-1.\end{aligned}$$

$$\text{综上有 } F_Z(z)=\begin{cases}0, & z\leqslant 48,\\ \frac{z}{32}-\frac{3}{2}, & 48<z<60,\\ \frac{9z}{160}-3, & 60\leqslant z<64,\\ \frac{z}{40}-1, & 64\leqslant z<80,\\ 1, & z\geqslant 80.\end{cases}$$

B12. 设一本书一页的错误个数 X 服从参数为 λ 的泊松分布, 且各页错误数相互独立. 现随机选 10 页, 其错误数分别记为 $X_1,X_2,\cdots,X_{10}$.

(1) 求 $P\left\{\sum_{i=1}^{10}X_i\geqslant 2\right\}$; (2) 求 $P\left\{\max_{1\leqslant i\leqslant 10}X_i\geqslant 2\right\}$;

(3) 求 $P\left\{\max_{1\leqslant i\leqslant 10}X_i\geqslant 2\,\middle|\,\min_{1\leqslant i\leqslant 10}X_i=0\right\}$.

解: (1) 事件 $\left\{\sum_{i=1}^{10}X_i\geqslant 2\right\}$ 的对立事件为 $\left\{\sum_{i=1}^{10}X_i<2\right\}=\left\{\sum_{i=1}^{10}X_i=0\right\}+\left\{\sum_{i=1}^{10}X_i=1\right\}$.

因为 $X_i\geqslant 0$, 则事件 $\left\{\sum_{i=1}^{10}X_i=0\right\}$ 表示 $X_i=0$, $i=1,2,\cdots 10$. 又因为各页错误相互独立有

$P\left\{\sum_{i=1}^{10}X_i=0\right\}=(P\{X_1=0\})^{10}=(e^{-\lambda})^{10}=e^{-10\lambda}$.

事件 $\left\{\sum_{i=1}^{10}X_i=1\right\}$ 表示 X_i, $i=1,2,\cdots 10$ 中有一个取值为 1, 剩下 9 个取值为 0.

$P\left\{\sum_{i=1}^{10}X_i=0\right\}=10(P\{X_1=0\})^9\cdot P\{X_1=1\}=10(e^{-\lambda})^9\cdot e^{-\lambda}\lambda=10\lambda e^{-10\lambda}.$

则 $P\left\{\sum_{i=1}^{10}X_i\leqslant 2\right\}=1-P\left\{\sum_{i=1}^{10}X_i=0\right\}-P\left\{\sum_{i=1}^{10}X_i=1\right\}=1-e^{-10\lambda}-10\lambda e^{-\lambda}.$

(2) 事件 $\left\{\max\limits_{1\leqslant i\leqslant 10}X_i\geqslant 2\right\}$ 的对立事件为 $\left\{\max\limits_{1\leqslant i\leqslant 10}X_i<2\right\}$.

也即意味着 $X_i\ (i=1,2,\cdots,10)$ 的取值小于 2, 即只能为 0 或 1.

$P\{X_1<2\}=P\{X_1=0\}+P\{X_1=1\}=e^{-\lambda}+\lambda e^{-\lambda}=(1+\lambda)e^{-\lambda}.$

从而 $P\left\{\max\limits_{1\leqslant i\leqslant 10}X_i\geqslant 2\right\}=1-P\left\{\max\limits_{1\leqslant i\leqslant 10}X_i<2\right\}=1-(P\{X_1<2\})^{10}=1-(1+\lambda)^{10}e^{-10\lambda}.$

(3) 事件 $\left\{\min\limits_{1\leqslant i\leqslant 10}X_i=0\right\}$ 的对立事件为 $\left\{\min\limits_{1\leqslant i\leqslant 10}X_i>0\right\}$.

即意味着 $X_i\ (i=1,2,\cdots,10)$ 的取值大于 0. 又因为 $P\{X_1>0\}=1-P\{X_1=0\}=1-e^{-\lambda}$.

从而 $P\left\{\min\limits_{1\leqslant i\leqslant 10}X_i=0\right\}=1-P\left\{\min\limits_{1\leqslant i\leqslant 10}X_i>0\right\}=1-(P\{X_1>0\})^{10}=1-(1-e^{-\lambda})^{10}.$

又因为事件

$\left\{\max\limits_{1\leqslant i\leqslant 10}X_i<2,\min\limits_{1\leq i\leq 10}X_i=0\right\}=\left\{\max\limits_{1\leqslant i\leqslant 10}X_i=1,\min\limits_{1\leqslant i\leqslant 10}X_i=0\right\}+\left\{\max\limits_{1\leqslant i\leqslant 10}X_i=0,\min\limits_{1\leqslant i\leqslant 10}X_i=0\right\}.$

$\left\{\max\limits_{1\leqslant i\leqslant 10}X_i=0,\min\limits_{1\leqslant i\leqslant 10}X_i=0\right\}$ 即表示 $X_i=0,\ i=1,2,\cdots,10$.

$$P\left\{\max_{1\leqslant i\leqslant 10}X_i=0,\min_{1\leqslant i\leqslant 10}X_i=0\right\}=(P\{X_1=0\})^{10}=(e^{-\lambda})^{10}.$$

$\left\{\max\limits_{1\leqslant i\leqslant 10}X_i=1,\min\limits_{1\leqslant i\leqslant 10}X_i=0\right\}$ 表示 $X_i\leqslant 1,\ i=1,2,\cdots,10$. 且至少有一个 1 和一个 0.

则只需刨去全是 0 和全是 1 的情况即可, 有

$$\begin{aligned}P\left\{\max_{1\leqslant i\leqslant 10}X_i=1,\min_{1\leqslant i\leqslant 10}X_i=0\right\}&=(P\{X_1\leqslant 1\})^{10}-(P\{X_1=1\})^{10}-(P\{X_1=0\})^{10}\\&=[(1+\lambda)e^{-\lambda}]^{10}-(e^{-\lambda})^{10}-(\lambda e^{-\lambda})^{10}.\end{aligned}$$

从而 $P\left\{\max\limits_{1\leqslant i\leqslant 10}X_i<2,\min\limits_{1\leq i\leq 10}X_i=0\right\}=[(1+\lambda)e^{-\lambda}]^{10}-(\lambda e^{-\lambda})^{10}=e^{-10\lambda}[(1+\lambda)^{10}-\lambda^{10}].$

因此

$$\begin{aligned}P\left\{\max_{1\leqslant i\leqslant 10}X_i\geqslant 2,\min_{1\leq i\leq 10}X_i=0\right\}&=P\left\{\min_{1\leq i\leq 10}X_i=0\right\}-P\left\{\max_{1\leqslant i\leqslant 10}X_i<2,\min_{1\leq i\leq 10}X_i=0\right\}\\&=1-(1-e^{-\lambda})^{10}-e^{-10\lambda}[(1+\lambda)^{10}-\lambda^{10}].\end{aligned}$$

从而

$$P\left\{\max_{1\leqslant i\leqslant 10} X_i \geqslant 2 \middle| \min_{1\leq i\leq 10} X_i = 0\right\} = \frac{P\left\{\max\limits_{1\leqslant i\leqslant 10} X_i \geqslant 2, \min\limits_{1\leq i\leq 10} X_i = 0\right\}}{P\left\{\min\limits_{1\leq i\leq 10} X_i = 0\right\}}$$

$$= \frac{1-(1-e^{-\lambda})^{10} - e^{-10\lambda}[(1+\lambda)^{10} - \lambda^{10}]}{1-(1-e^{-\lambda})^{10}}$$

$$= 1 - \frac{e^{-10\lambda}[(1+\lambda)^{10} - \lambda^{10}]}{1-(1-e^{-\lambda})^{10}}.$$

B13. 设一系统由两个独立的子系统组成, 分别以 X, Y 记两个子系统的正常工作时间, 且设 X, Y 分别服从参数为 λ_1 与 λ_2 的指数分布. 当这两个子系统 (1) 串联, (2) 并联, (3) 有备份 (当一个损坏时另一个接着工作) 时, 分别求系统正常工作时间 T 的密度函数.

解: 由题 X 与 Y 的分布函数分别为 $F_X(x) = \begin{cases} 1-e^{-\lambda_1 x}, & x>0, \\ 0, & x\leqslant 0. \end{cases}$ $F_Y(y) = \begin{cases} 1-e^{-\lambda_2 y}, & y>0, \\ 0, & y\leqslant 0. \end{cases}$

(1) 串联工作时, 只要有一个子系统损坏, 系统便损坏.

因此系统正常工作的时间取决于两个子系统中工作时间较短的那个.

设此时系统正常工作时间为 T_1, 则 $T_1 = \min(X, Y)$.

从而 $t>0$ 时 T_1 的分布函数 $F_{T_1}(t) = 1-[1-F_X(t)]\cdot[1-F_Y(t)] = 1-\mathrm{e}^{-\lambda_1 t}\cdot\mathrm{e}^{-\lambda_2 t} = 1-\mathrm{e}^{-(\lambda_1+\lambda_2)t}$.

此时 T_1 的密度函数 $f_{T_1}(t) = F'_{T_1}(t) = -[-(\lambda_1+\lambda_2)]\mathrm{e}^{-(\lambda_1+\lambda_2)t} = (\lambda_1+\lambda_2)\mathrm{e}^{-(\lambda_1+\lambda_2)t}$.

又 $t\leqslant 0$ 时 $F_{T_1}(t) = 1-(1-0)\cdot(1-0) = 0$, 则 $f_{T_1}(t) = 0$. 从而

$$f_{T_1}(t) = \begin{cases} (\lambda_1+\lambda_2)\mathrm{e}^{-(\lambda_1+\lambda_2)t}, & t>0, \\ 0, & t\leqslant 0. \end{cases}$$

(2) 并联工作时, 两个子系统都损坏时, 系统才损坏.

因此系统正常工作的时间取决于两个子系统中工作时间较长的那个.

设此时系统正常工作时间为 T_2, 则 $T_2 = \max(X, Y)$.

从而 $t>0$ 时 T_2 的分布函数

$$F_{T_2}(t) = F_X(t)\cdot F_Y(t) = (1-\mathrm{e}^{-\lambda_1 t})\cdot(1-\mathrm{e}^{-\lambda_2 t}) = 1-(\mathrm{e}^{-\lambda_1 t}+\mathrm{e}^{-\lambda_2 t}) + \mathrm{e}^{-(\lambda_1+\lambda_2)t}.$$

此时 T_2 的密度函数 $f_{T_2}(t) = F'_{T_2}(t) = (\lambda_1 e^{-\lambda_1 t} + \lambda_2 e^{-\lambda_2 t}) - (\lambda_1+\lambda_2)\mathrm{e}^{-(\lambda_1+\lambda_2)t}$.

又 $t \leqslant 0$ 时 $F_{T_2}(t) = 0 \cdot 0 = 0$, 则 $f_{T_2}(t) = 0$. 从而

$$f_{T_2}(t) = \begin{cases} (\lambda_1 e^{-\lambda_1 t} + \lambda_2 e^{-\lambda_2 t}) - (\lambda_1 + \lambda_2)\mathrm{e}^{-(\lambda_1+\lambda_2)t}, & t > 0, \\ 0, & t \leqslant 0. \end{cases}$$

(3) 有备份时, 设备份子系统工作时间为为 Y, 从而系统正常工作时间 $T_3 = X + Y$.

从而 T_3 的密度函数 $f_{T_3}(t) = \int_{-\infty}^{+\infty} f_X(x) \cdot f_Y(t-x)\mathrm{d}x$.

则 $t \leqslant 0$ 时, $f_{T_3}(t) = 0$.

$t > 0$ 时, $f_X(x) \cdot f_Y(t-x) = \lambda_1 \mathrm{e}^{-\lambda_1 x} \cdot \lambda_2 \mathrm{e}^{-\lambda_2 (t-x)} = \lambda_1\lambda_2 \mathrm{e}^{(\lambda_2-\lambda_1)x}\mathrm{e}^{-\lambda_2 t}$, 此时 $0 < x < t$.

当 $\lambda_1 \neq \lambda_2$ 时,

$$\begin{aligned} f_{T_3}(t) &= \int_{-\infty}^{+\infty} f_X(x) \cdot f_Y(t-x)\mathrm{d}x = \int_0^t \lambda_1\lambda_2 \mathrm{e}^{(\lambda_2-\lambda_1)x}\mathrm{e}^{-\lambda_2 t}\mathrm{d}x \\ &= \lambda_1\lambda_2\mathrm{e}^{-\lambda_2 t} \cdot \frac{1}{\lambda_2-\lambda_1}\mathrm{e}^{(\lambda_2-\lambda_1)x}\Big|_0^t = \frac{\lambda_1\lambda_2}{\lambda_2-\lambda_1}\mathrm{e}^{-\lambda_2 t}\left[\mathrm{e}^{(\lambda_2-\lambda_1)t} - 1\right] \\ &= \frac{\lambda_1\lambda_2}{\lambda_2-\lambda_1}(\mathrm{e}^{-\lambda_1 t} - \mathrm{e}^{-\lambda_2 t}). \end{aligned}$$

因此 $f_{T_3}(t) = \begin{cases} \dfrac{\lambda_1\lambda_2}{\lambda_2-\lambda_1}(\mathrm{e}^{-\lambda_1 t} - \mathrm{e}^{-\lambda_2 t}), & t > 0, \\ 0, & t \leq 0. \end{cases}$

当 $\lambda_1 = \lambda_2 = \lambda$ 时, 对 $t > 0$,

$$\begin{aligned} f_{T_3}(t) &= \int_{-\infty}^{+\infty} f_X(x) \cdot f_Y(t-x)\mathrm{d}x = \int_0^t \lambda_1\lambda_2 \mathrm{e}^{(\lambda_2-\lambda_1)x}\mathrm{e}^{-\lambda_2 t}\mathrm{d}x \\ &= \lambda^2\mathrm{e}^{-\lambda t}\int_0^t 1\mathrm{d}x = \lambda^2 t\mathrm{e}^{-\lambda t}. \end{aligned}$$

因此 $f_{T_3}(t) = \begin{cases} \lambda^2 t\mathrm{e}^{-\lambda t}, & t > 0, \\ 0, & t \leq 0. \end{cases}$

第 4 章　随机变量的数字特征

A1. 设一盒中有 3 个红球 2 个白球. 从中取 3 次, 每次取 1 个球. 令 X 表示取到红球的个数, 分别求 (1)无放回抽样, (2) 有放回抽样这两种抽样方式下的 $E(X)$.

解: (1) 无放回时, 不可能取出的 3 个球中没有红球, 则 X 可能取 $1,2,3$.

$P\{X=1\}=\frac{2}{5}\cdot\frac{1}{4}\cdot\frac{3}{3}+\frac{3}{5}\cdot\frac{2}{4}\cdot\frac{1}{3}+\frac{2}{5}\cdot\frac{3}{4}\cdot\frac{1}{3}=\frac{6+6+6}{60}=\frac{3}{10}$.

$P\{X=2\}=\frac{2}{5}\cdot\frac{3}{4}\cdot\frac{2}{3}+\frac{3}{5}\cdot\frac{2}{4}\cdot\frac{2}{3}+\frac{3}{5}\cdot\frac{2}{4}\cdot\frac{2}{3}=\frac{12+12+12}{60}=\frac{3}{5}$.

$P\{X=3\}=\frac{3}{5}\cdot\frac{2}{4}\cdot\frac{1}{3}=\frac{6}{60}=\frac{1}{10}$.

则 $E(X)=\sum\limits_{k=0}^{3}k\cdot P\{X=k\}=0+\frac{3}{10}+2\times\frac{3}{5}+3\times\frac{1}{10}=\frac{9}{5}$.

(2) 有放回时, 每次抽中红球概率为 $\frac{3}{5}$, 每次抽中白球概率为 $\frac{2}{5}$.

$P\{X=0\}=\frac{2}{5}\cdot\frac{2}{5}\cdot\frac{2}{5}=\frac{8}{125}$; $P\{X=1\}=C_3^1\cdot\frac{3}{5}\cdot\frac{2}{5}\cdot\frac{2}{5}=\frac{36}{125}$;

$P\{X=2\}=C_3^2\cdot\frac{3}{5}\cdot\frac{3}{5}\cdot\frac{2}{5}=\frac{54}{125}$; $P\{X=3\}=\frac{3}{5}\cdot\frac{3}{5}\cdot\frac{3}{5}=\frac{27}{125}$.

则 $E(X)=\sum\limits_{k=0}^{3}k\cdot P\{X=k\}=0+\frac{36}{125}+2\times\frac{54}{125}+3\times\frac{27}{125}=\frac{225}{125}=\frac{9}{5}$.

A2. 设随机变量 X 的密度函数为

$$f(x)=\begin{cases}\frac{x}{4}, & 1<x<3,\\ 0, & \text{其他},\end{cases}$$

求 $E(X)$ 和 $P\{X>E(X)\}$.

解:

$$E(X)=\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x=\int_1^3\frac{x^2}{4}\mathrm{d}x=\left.\frac{x^3}{12}\right|_1^3=\frac{27}{12}-\frac{1}{12}=\frac{13}{6}.$$

$$\begin{aligned}P\{X>E(X)\}&=P\left\{X>\frac{13}{6}\right\}=1-P\left\{X\leqslant\frac{13}{6}\right\}=1-\int_{-\infty}^{\frac{13}{6}}f(x)\mathrm{d}x=1-\int_1^{\frac{13}{6}}\frac{x}{4}\mathrm{d}x\\&=1-\left.\frac{x^2}{8}\right|_1^{\frac{13}{6}}=1-\left(\frac{1}{8}\cdot\frac{13^2}{6^2}-\frac{1}{8}\right)=1-\frac{169-36}{288}=\frac{155}{288}.\end{aligned}$$

A3. 设随机变量 X 的密度函数为

$$f(x)=\begin{cases} ax+b, & 0<x<2, \\ 0, & \text{其他}, \end{cases}$$

已知 $E(X)=\dfrac{11}{9}$，求 a 和 b.

解：由题，$1=\displaystyle\int_{-\infty}^{+\infty} f(x)\mathrm{d}x=\int_0^2 (ax+b)\mathrm{d}x=\left(\frac{1}{2}ax^2+bx\right)\Big|_0^2=2a+2b.$

且 $E(X)=\displaystyle\int_{-\infty}^{+\infty} xf(x)\mathrm{d}x=\int_0^2 x(ax+b)\mathrm{d}x=\left(\frac{1}{3}ax^3+\frac{1}{2}bx^2\right)\Big|_0^2=\frac{8a}{3}+2b.$

由题 $E(X)=\dfrac{11}{9}$ 联立有 $\begin{cases} 2a+2b=1, \\ \dfrac{8}{3}a+2b=\dfrac{11}{9} \end{cases}$ 解得 $\begin{cases} a=\dfrac{1}{3}, \\ b=\dfrac{1}{6}. \end{cases}$

A4. 设二维离散型随机变量 (X,Y) 的联合分布律为

X \ Y	0	1	2
1	0.2	0.1	0.3
2	0.2	0	0.2

求随机变量 Z 的数学期望 $E(Z)$:

(1) $Z=XY$; (2) $Z=\min\{X,Y\}$; (3) $Z=\max\{X,Y\}$.

解：(1) 此时 Z 可能取值为 $0,1,2,4$.

$P\{Z=0\}=P\{Y=0\}=0.4$; $P\{Z=1\}=P\{X=Y=1\}=0.1$;

$P\{Z=2\}=P\{X=1,Y=2\}+P\{X=2,Y=1\}=0.3$; $P\{Z=4\}=P\{X=Y=2\}=0.2$.

则其概率分布律为

Z	0	1	2	4
p	0.4	0.1	0.3	0.2

从而 $E(Z)=0\times0.4+1\times0.1+2\times0.3+4\times0.2=1.5.$

(2) 此时 Z 可能取值为 $0,1,2$.

$P\{Z=0\}=P\{Y=0\}=0.2+0.2=0.4$; $P\{Z=2\}=P\{X=Y=2\}=0.2$.

则 $P\{Z=1\}=1-0.4-0.2=0.4$，从而其概率分布律为

Z	0	1	2
p	0.4	0.4	0.2

从而 $E(Z)=0\times 0.4+1\times 0.4+2\times 0.2=0.8.$

(3) 此时 Z 可能取值为 $1,2.$

$P\{Z=1\}=P\{X=1,Y=0\}+P\{X=1,Y=1\}=0.3;\ P\{Z=2\}=1-0.3=0.7.$

其概率分布律为

Z	1	2
p	0.3	0.7

从而 $E(Z)=1\times 0.3+2\times 0.7=1.7.$

A5. 设二维连续型随机变量 (X,Y) 的联合密度函数为

$$f(x,y)=\begin{cases}\dfrac{2x+y}{4}, & 0<x<1,0<y<2,\\ 0, & \text{其他},\end{cases}$$

分别求数学期望 $E(X)$, $E(Y)$, $E(XY)$.

解: X 的边际密度函数

$$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}y=\int_0^2\frac{2x+y}{4}\mathrm{d}y=\left(\frac{xy}{2}+\frac{y^2}{8}\right)\Big|_0^2=x+\frac{1}{2},\ 0<x<1.$$

从而

$$E(X)=\int_0^1 xf_X(x)\mathrm{d}x=\int_0^1\left(x^2+\frac{x}{2}\right)\mathrm{d}x=\left(\frac{x^3}{6}+\frac{x^2}{4}\right)\Big|_0^1=\frac{7}{12}.$$

Y 的边际密度函数

$$f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x=\int_0^1\frac{2x+y}{4}\mathrm{d}x=\left(\frac{x^2}{4}+\frac{xy}{4}\right)\Big|_0^1=\frac{y}{4}+\frac{1}{4},\ 0<y<2.$$

从而

$$E(Y)=\int_0^2 yf_Y(y)\mathrm{d}y=\int_0^2\left(\frac{y^2}{4}+\frac{y}{4}\right)\mathrm{d}y=\left(\frac{y^3}{12}+\frac{y^2}{8}\right)\Big|_0^2=\frac{8}{12}+\frac{4}{8}=\frac{7}{6}.$$

$$\begin{aligned}E(XY)&=\int_{-\infty}^{+\infty}\int_{\infty}^{+\infty}xyf(x,y)\mathrm{d}x=\int_0^2\int_0^1 xy\cdot\frac{2x+y}{4}\mathrm{d}x\mathrm{d}y\\&=\int_0^2\int_0^1\left(\frac{yx^2}{2}+\frac{y^2x}{4}\right)\mathrm{d}x\mathrm{d}y=\int_0^2\left(\frac{1}{6}yx^3+\frac{1}{8}y^2x^2\right)\Big|_0^1\mathrm{d}y\\&=\int_0^2\left(\frac{1}{6}y+\frac{1}{8}y^2\right)\mathrm{d}y=\left(\frac{1}{12}y^2+\frac{1}{24}y^3\right)\Big|_0^2\\&=\frac{2^2}{12}+\frac{2^3}{24}=\frac{2}{3}.\end{aligned}$$

A6. 设随机变量 X 的密度函数为

$$f(x)=\begin{cases}3x^2, & 0<x<1,\\ 0, & \text{其他},\end{cases}$$

求 $E\left(3X^2-\dfrac{1}{3X^2}\right)$.

解: 由于

$$\begin{aligned}\int_{-\infty}^{+\infty}\left(3x^2-\frac{1}{3x^2}\right)f(x)\mathrm{d}x&=\int_0^1\left(3x^2-\frac{1}{3x^2}\right)\cdot 3x^2\mathrm{d}x=\int_0^1(9x^4-1)\mathrm{d}x\\&=\left(\frac{9}{5}x^5-x\right)\Big|_0^1=\frac{9}{5}-1=\frac{4}{5},\end{aligned}$$

则 $E\left(3X^2-\dfrac{1}{3X^2}\right)=\dfrac{4}{5}$.

A7. 设一盒中有 3 个红球 2 个白球, 从中取 3 次, 每次取 1 个球, 令 X 表示取到红球的个数, 分别求 (1) 无放回抽样, (2) 有放回抽样两种抽样方式下的 $\mathrm{Var}(X)$.

解: (1) 由 A1 可知, $E(X)=\dfrac{9}{5}$. 从而有

$$\begin{aligned}\mathrm{Var}(X)&=(1-E(X))^2\cdot P\{X=1\}+(2-E(X))^2\cdot P\{X=2\}+(3-E(X))^2\cdot P\{X=3\}\\&=\frac{16}{25}\cdot\frac{3}{10}+\frac{1}{25}\cdot\frac{3}{5}+\frac{36}{25}\cdot\frac{1}{10}=\frac{24+3+18}{125}\\&=\frac{45}{125}=\frac{9}{25}.\end{aligned}$$

(2) 由 A1 可知, $E(X)=\dfrac{9}{5}$, 且

$E(X^2)=\displaystyle\sum_{k=0}^{4}k^2\cdot P\{X=k\}=0+\frac{36}{125}+\frac{2^2\times 54}{125}+\frac{3^2\times 27}{125}=\frac{36+216+243}{125}=\frac{495}{125}=\frac{99}{25}.$

从而 $\mathrm{Var}(X)=E(X^2)-(E(X))^2=\dfrac{99}{25}-\dfrac{81}{25}=\dfrac{18}{25}$.

A8. 设随机变量 X 的密度函数为

$$f(x)=\begin{cases}\dfrac{3x^2}{2}, & -1<x<1,\\ 0, & \text{其他},\end{cases}$$

求 $\mathrm{Var}(X)$.

解: 由题有 $E(X)=\displaystyle\int_{-1}^{1}x\cdot\frac{3x^2}{2}\mathrm{d}x=0$(奇函数的对称性); $E(X^2)=\displaystyle\int_{-1}^{1}x^2\cdot\frac{3x^2}{2}\mathrm{d}x=\frac{3x^5}{10}\Big|_{-1}^{1}=\frac{3}{5}$.

因此 $\mathrm{Var}(X)=E(X^2)-(E(X))^2=\dfrac{3}{5}$.

A9. 设随机变量 X 与 Y 相互独立, 密度函数分别为

$$f_X(x)=\begin{cases}\dfrac{3x^2}{2}, & -1<x<1,\\ 0, & \text{其他},\end{cases}\qquad f_Y(y)=\begin{cases}|y|, & -1<y<1,\\ 0, & \text{其他},\end{cases}$$

求 $\mathrm{Var}(X-2Y)$.

解: 由 A8 可知 $\mathrm{Var}(X)=\dfrac{3}{5}$.

$E(Y)=\displaystyle\int_{-1}^{1}y|y|\mathrm{d}y=0$ (奇函数); $E(Y^2)=\displaystyle\int_{-1}^{1}y^2|y|\mathrm{d}y=2\int_0^1 y^3\mathrm{d}y=2\cdot\frac{1}{4}=\frac{1}{2}$(对称性).

则 $\mathrm{Var}(Y)=E(Y^2)-(E(Y))^2=\dfrac{1}{2}$. 因此 $\mathrm{Var}(2Y)=4\mathrm{Var}(Y)=2$.

又因为 X,Y 相互独立, 所以 $\mathrm{Var}(X-2Y)=\mathrm{Var}(X)+\mathrm{Var}(2Y)=\dfrac{13}{5}$.

A10. 设 X 的分布函数如下:

$$(1)\ F(x)=\begin{cases}0, & x<0,\\ 0.3, & 0\leqslant x<1,\\ 0.5, & 1\leqslant x<2,\\ 1, & x\geqslant 2,\end{cases}$$

$$(2)\ F(x)=\begin{cases}0, & x<0,\\ \dfrac{x^2}{2}, & 0\leqslant x<1,\\ \dfrac{x}{2}, & 1\leqslant x<2,\\ 1, & x\geqslant 2.\end{cases}$$

分别求 $E(X)$, $\mathrm{Var}(X)$.

解: (1) 由分布函数可得 X 的概率分布律为

X	0	1	2
p	0.3	0.2	0.5

从而 $E(X)=0\times0.3+1\times0.2+2\times0.5=1.2$.

$\mathrm{Var}(X)=(0-1.2)^2\times0.3+(1-1.2)^2\times0.2+(2-1.2)^2\times0.5=0.432+0.008+0.32=0.76$.

(2) 由分布函数, X 的密度函数 $f_X(x)=\begin{cases}x, & 0\leqslant x<1,\\ \dfrac{1}{2}, & 1\leqslant x<2,\\ 0, & 其他.\end{cases}$

从而 $E(X)=\int_{-\infty}^{+\infty}xf_X(x)\mathrm{d}x=\int_0^1 x^2\mathrm{d}x+\int_1^2\frac{x}{2}\mathrm{d}x=\frac{x^3}{3}\Big|_0^1+\frac{x^2}{4}\Big|_1^2=\frac{1}{3}+1-\frac{1}{4}=\frac{13}{12}$.

$E(X^2)=\int_{-\infty}^{+\infty}x^2f_X(x)\mathrm{d}x=\int_0^1 x^3\mathrm{d}x+\int_1^2\frac{x^2}{2}\mathrm{d}x=\frac{x^4}{4}\Big|_0^1+\frac{x^3}{6}\Big|_1^2=\frac{1}{4}+\frac{8}{6}-\frac{1}{6}=\frac{3}{12}+\frac{14}{12}=\frac{17}{12}$.

因此 $\mathrm{Var}(X)=E(X^2)-(E(X))^2=\frac{17}{12}-\frac{169}{144}=\frac{204-169}{144}=\frac{35}{144}$.

A11. 设 (X,Y) 在区域 $D=\{(x,y)\mid 0<x<y<1\}$ 上服从均匀分布, 求 $\mathrm{Var}(XY)$.

解: 区域 D 为点 $(0,0)$, $(1,1)$, $(0,1)$ 所围三角形的内部，其面积为 $\frac{1}{2}$.

从而 $f_{XY}(x,y)=\begin{cases}2, & (x,y)\in D,\\ 0, & 其他.\end{cases}$

$E(XY)=\int_0^1\int_0^y 2xy\mathrm{d}x\mathrm{d}y=\int_0^1 x^2y\Big|_0^y\mathrm{d}y=\int_0^1 y^3\mathrm{d}y=\frac{1}{4}$.

$E(X^2Y^2)=\int_0^1\int_0^y 2x^2y^2\mathrm{d}x\mathrm{d}y=\int_0^1\frac{2x^3y^2}{3}\Big|_0^y\mathrm{d}y=\int_0^1\frac{2y^5}{3}\mathrm{d}y=\frac{2y^6}{18}\Big|_0^1=\frac{1}{9}$.

从而 $\mathrm{Var}(XY)=E(X^2Y^2)-(E(XY))^2=\frac{1}{9}-\frac{1}{16}=\frac{7}{144}$.

A12. 设随机变量 X 与 Y 相互独立, $X\sim P(2)$, $Y\sim B(2,0.4)$, 求 $E(2X-Y)$, $\mathrm{Var}(2X-Y)$, $E[(2X-Y)^2]$.

解: 由泊松分布, $E(X)=2$, $\mathrm{Var}(X)=2$.

由二项分布, $E(Y)=2\times0.4=0.8$, $\mathrm{Var}(Y)=2\times0.4\times(1-0.4)=0.48$.

所以 $E(2X-Y)=2E(X)-E(Y)=3.2$.

又因为 X,Y 相互独立, 从而 $\mathrm{Var}(2X-Y)=\mathrm{Var}(2X)+\mathrm{Var}(Y)=4\mathrm{Var}(X)+0.48=8.48$.

$E[(2X-Y)^2]=\mathrm{Var}(2X-Y)+(E(2X-Y))^2=8.48+3.2^2=18.72$.

A13. 设随机变量 X 服从参数为 $\frac{1}{2}$ 的指数分布, 求 $E[X(X-1)]$.

解: 由题, $E(X)=2$, $\mathrm{Var}(X)=2^2=4=E(X^2)-(E(X))^2$.

则 $E(X^2)=8$, 从而 $E[X(X-1)]=E(X^2-X)=E(X^2)-E(X)=8-2=6$.

A14. 设随机变量 $X\sim U(a,b)$, $Y\sim N(2,3)$, X 与 Y 有相同的数学期望与方差, 求 a, b 的值.

解: 由题 $E(Y)=2$, $\mathrm{Var}(Y)=3$.

而 $E(X)=\int_a^b \frac{x}{b-a}\mathrm{d}x=\left.\frac{x^2}{2(b-a)}\right|_a^b=\frac{b^2-a^2}{2(b-a)}=\frac{b+a}{2}$.

$E(X^2)=\int_a^b \frac{x^2}{b-a}\mathrm{d}x=\left.\frac{x^3}{3(b-a)}\right|_a^b=\frac{b^3-a^3}{3(b-a)}=\frac{b^2+ab+a^2}{3}$.

$$\begin{aligned}\mathrm{Var}(X)=E(X^2)-(E(X))^2&=\frac{b^2+ab+a^2}{3}-\frac{(b+a)^2}{4}\\&=\frac{4(b^2+ab+a^2)-3(b^2+2ab+a^2)}{12}\\&=\frac{b^2-2ab+a^2}{12}=\frac{(b-a)^2}{12}.\end{aligned}$$

从而有 $\frac{b+a}{2}=2$, $\frac{(b-a)^2}{12}=3$. 由于 $b>a$ 也即 $\begin{cases}a+b=4,\\ b-a=6.\end{cases}$

解得 $a=-1$, $b=5$.

A15. 设二维离散型随机变量 (X,Y) 的联合分布律如第 A4 题所示, 求协方差 $\mathrm{Cov}(X,Y)$ 及相关系数 ρ_{XY}.

解: 由 A4 可知, $P\{X=1\}=0.6$, $P\{X=2\}=0.4$.

则 $E(X)=1\times 0.6+2\times 0.4=1.4$; $\mathrm{Var}(X)=(1-1.4)^2\times 0.6+(2-1.4)^2\times 0.4=0.096+0.144=0.24$.

又因为 $P\{Y=0\}=0.4$, $P\{Y=1\}=0.1$, $P\{Y=2\}=0.5$.

从而 $E(Y)=0\times 0.4+1\times 0.1+2\times 0.5=1.1$;

$\mathrm{Var}(Y)=(0-1.1)^2\times 0.4+(1-1.1)^2\times 0.1+(2-1.1)^2\times 0.5=0.484+0.001+0.405=0.89$.

由 A4(1) 可知, $E(XY)=1.5$, 从而 $\mathrm{Cov}(X,Y)=E(XY)-E(X)E(Y)=1.5-1.4\times 1.1=-0.04$.

$\rho_{XY}=\frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}=\frac{-0.04}{\sqrt{0.24\times 0.89}}=-\frac{0.04}{0.02\sqrt{6\times 89}}=-\frac{2}{\sqrt{534}}=-\frac{\sqrt{534}}{267}$.

A16. 设随机变量 (X,Y) 的联合密度函数为

$$f(x,y)=\begin{cases}\frac{3x}{4}, & 0<x<y<2,\\ 0, & \text{其他},\end{cases}$$

求协方差 $\mathrm{Cov}(X,Y)$ 及相关系数 ρ_{XY}.

解: 由题

$$E(X)=\int_0^2\mathrm{d}x\int_x^2 xf(x,y)\mathrm{d}y=\int_0^2\mathrm{d}x\int_x^2\frac{3}{4}x^2\mathrm{d}y=\int_0^2\left.\frac{3}{4}x^2y\right|_x^2\mathrm{d}x=\frac{3}{4}\int_0^2 x^2(2-x)\mathrm{d}x$$

$$= \frac{3}{4}\left(-\frac{1}{4}x^4 + \frac{2}{3}x^3\right)\Big|_0^2 = \frac{3}{4}\cdot\left(-\frac{2^4}{4} + \frac{2}{3}\cdot 2^3\right) = \frac{3}{4}\cdot\left(\frac{16}{3} - 4\right) = 1.$$

$$\begin{aligned} E(X^2) &= \int_0^2 \mathrm{d}x \int_x^2 x^2 f(x,y)\mathrm{d}y = \int_0^2 \mathrm{d}x \int_x^2 \frac{3}{4}x^3 \mathrm{d}y = \int_0^2 \frac{3}{4}x^3 y\Big|_x^2 \mathrm{d}x = \frac{3}{4}\int_0^2 x^3(2-x)\mathrm{d}x \\ &= \frac{3}{4}\left(-\frac{1}{5}x^5 + \frac{1}{2}x^4\right)\Big|_0^2 = \frac{3}{4}\cdot\left(-\frac{2^5}{5} + \frac{1}{2}\cdot 2^4\right) = \frac{3}{4}\cdot\left(8 - \frac{32}{5}\right) = \frac{6}{5}. \end{aligned}$$

$$\begin{aligned} E(Y) &= \int_0^2 \mathrm{d}y \int_0^y y f(x,y)\mathrm{d}x = \int_0^2 \mathrm{d}y \int_0^y \frac{3}{4}xy \mathrm{d}x = \int_0^2 \frac{3}{8}yx^2\Big|_0^y \mathrm{d}y = \frac{3}{8}\int_0^2 y(y^2-0)\mathrm{d}y \\ &= \frac{3}{8}\cdot\frac{1}{4}y^4\Big|_0^2 = \frac{3}{32}\cdot 2^4 = \frac{3}{2}. \end{aligned}$$

$$\begin{aligned} E(Y^2) &= \int_0^2 \mathrm{d}y \int_0^y y^2 f(x,y)\mathrm{d}x = \int_0^2 \mathrm{d}y \int_0^y \frac{3}{4}xy^2 \mathrm{d}x = \int_0^2 \frac{3}{8}y^2x^2\Big|_0^y \mathrm{d}y = \frac{3}{8}\int_0^2 y^2(y^2-0)\mathrm{d}y \\ &= \frac{3}{8}\cdot\frac{1}{5}y^5\Big|_0^2 = \frac{3}{40}\cdot 2^5 = \frac{12}{5}. \end{aligned}$$

所以 $\mathrm{Var}(X) = E(X^2) - (E(X))^2 = \frac{6}{5} - 1^2 = \frac{1}{5} = 0.2$;

$\mathrm{Var}(Y) = E(Y^2) - (E(Y))^2 = \frac{12}{5} - \frac{3^2}{2^2} = \frac{48}{20} - \frac{45}{20} = \frac{3}{20} = 0.15.$

$$\begin{aligned} E(XY) &= \int_0^2\int_0^y xyf(x,y)\mathrm{d}x\mathrm{d}y = \int_0^2\int_0^y \frac{3x^2y}{4}\mathrm{d}x\mathrm{d}y = \int_0^2 \frac{x^3y}{4}\Big|_0^y \mathrm{d}y \\ &= \frac{1}{4}\int_0^2 y^4\mathrm{d}y = \frac{1}{4}\cdot\frac{1}{5}y^5\Big|_0^2 = \frac{1}{20}(2^5-0) = \frac{8}{5}. \end{aligned}$$

从而 $\mathrm{Cov}(X,Y) = E(XY) - E(X)E(Y) = \frac{8}{5} - 1\cdot\frac{3}{2} = \frac{1}{10} = 0.1.$

$\rho_{XY} = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}} = \frac{0.1}{\sqrt{0.2\times 0.15}} = \frac{0.1}{\sqrt{0.03}} = \frac{1}{\sqrt{3}} = \frac{\sqrt{3}}{3}.$

A17. 设 $(X,Y)\sim N(-1,1;4,4;0.6)$, 求 $E(XY)$.

解: 此时 $E(X) = -1$, $\mathrm{Var}(X) = 4$; $E(Y) = 1$, $\mathrm{Var}(Y) = 4$; $\rho_{XY} = 0.6$.

从而 $\mathrm{Cov}(X,Y) = \rho_{XY}\cdot\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)} = 0.6\times\sqrt{4\times 4} = 2.4.$

则 $E(XY) = E(X)E(Y) + \mathrm{Cov}(X,Y) = -1 + 2.4 = 1.4.$

A18. 设二维离散型随机变量 (X,Y) 的联合分布律为

X	Y		
	0	1	2
1	0.1	0.4	0.1
2	0.2	0	0.2

判断 X 与 Y 是否相关, 是否相互独立.

解: 由题对 X, $P\{X=1\}=0.6$, $P\{X=2\}=0.4$.

从而 $E(X)=1\times 0.6+2\times 0.4=1.4$.

对 Y, $P\{Y=0\}=0.3$, $P\{Y=1\}=0.4$, $P\{Y=2\}=0.3$.

从而 $E(Y)=0\times 0.3+1\times 0.4+2\times 0.3=1$.

由联合分布律, XY 可能取值为 $0,1,2,4$, 其概率分布律为

XY	0	1	2	4
p	0.3	0.4	0.1	0.2

则 $E(XY)=0\times 0.3+1\times 0.4+2\times 0.1+4\times 0.2=1.4$.

所以有 $E(XY)=E(X)E(Y)$ 成立.

因此 X 与 Y 不相关.

又因为 $P\{X=1,Y=0\}=0.1$, $P\{X=1\}\cdot P\{Y=0\}=0.6\times 0.3=0.18$.

此时 $P\{X=1,Y=0\}\neq P\{X=1\}\cdot P\{Y=0\}$, 则 X 与 Y 不独立.

A19. 设 $(X,Y)\sim N(-1,1;4,4;0.6)$, 若 $X+Y$ 与 $X-aY$ 相互独立, 此时 a 取何值?

解: 由二维正态分布可知, 若 $X+Y$ 与 $X-aY$ 相互独立, 则二者不相关, 即 $\mathrm{Cov}(X+Y,X-aY)=0$.

因为 $\mathrm{Cov}(X+Y,X-aY)=\mathrm{Cov}(X,X)-a\mathrm{Cov}(Y,Y)+(1-a)\mathrm{Cov}(X,Y)=$

$\mathrm{Var}(X)-a\mathrm{Var}(Y)+(1-a)\mathrm{Cov}(X,Y)=0$.

且由题和 A17 的结论, $\mathrm{Var}(X)=4$, $\mathrm{Var}(Y)=4$, $\mathrm{Cov}(X,Y)=2.4$.

从而 $4-4a+2.4(1-a)=0$, 解得 $a=1$.

B1. 某批产品共有 M 件, 其中正品 N 件 $(0 \leqslant N \leqslant M)$. 从整批产品中随机地进行有放回抽样: 每次抽取一件, 记录产品是正品还是次品后放回, 抽取了 n 次 $(n \geqslant 1)$, 求这 n 次中抽到正品的平均次数.

解: 由题, 设抽 n 次中抽到正品次数为 X.

由于每次抽中正品的概率为 $\dfrac{N}{M}$, 从而 $X \sim B\left(n, \dfrac{N}{M}\right)$.

平均次数即为 $E(X) = n \cdot \dfrac{N}{M} = \dfrac{nN}{M}$.

B2. 一位即将毕业的大学生有意向与某企业签订就业合同. 该企业给他两个年薪方案供选择: 方案一: 年薪 10 万元; 方案二: 底薪 6 万元, 如果业绩达到公司要求, 就可再获得业绩津贴 10 万元, 如果达不到, 就没有业绩津贴, 一般约有 80% 的可能性可以达到公司的业绩要求. 他应当选择哪种方案? 说明理由.

解: 记方案二的年薪为 X 万元, 则 X 可以取 6, 16.

且 $P\{X = 6\} = 0.2$, $P\{X = 16\} = 0.8$, 从而 $E(X) = 6 \times 0.2 + 16 \times 0.8 = 14$.

因为 $14 > 10$, 则应该选方案二, 因为平均年薪较高.

B3. 设一袋中有 8 个球, 分别编号为 $1 \sim 8$. 现随机从袋中取出 2 球, 记其中最大号码的球的编号为 X, 求 $E(X)$.

解: 总取球次数为 $\mathrm{C}_8^2 = \dfrac{8 \times 7}{2} = 28$. 此时 X 可能取值为 $2, 3, 4, 5, 6, 7, 8$.

$X = i$ 时, 对应取出 i 号球, 另一球数字比 i 小即可, 所以此时情况数为 $i - 1$ 种.

即 $X = 2$ 时有一种情况, $X = 3$ 时有两种情况, $\cdots$, $X = 8$ 时有七种情况.

从而 X 的概率分布律为

X	2	3	4	5	6	7	8
P	$\frac{1}{28}$	$\frac{2}{28}$	$\frac{3}{28}$	$\frac{4}{28}$	$\frac{5}{28}$	$\frac{6}{28}$	$\frac{7}{28}$

$$\begin{aligned} E(X) &= \sum_{i=2}^{8} i \cdot P\{X = i\} = \frac{2}{28} + \frac{6}{28} + \frac{12}{28} + \frac{20}{28} + \frac{30}{28} + \frac{42}{28} + \frac{56}{28} \\ &= \frac{168}{28} = 6. \end{aligned}$$

B4. 直线上一质点在时刻 0 从原点出发, 每经过一个单位时间向左或者向右移动一个单位, 每次移动是相互独立的, 并且向右移动的概率为 p $(0 < p < 1)$. 记 η_n 表示到时刻 n 为止质点向右移动的次数, S_n 表示在时刻 n 时质点的位置 $(n \geqslant 1)$, 求 η_n 与 S_n 的数学期望.

解: 由题有 $\eta_n \sim B(n, p)$, 故 $E(\eta_n) = np$.

又因为此时 $n - \eta_n$ 表示向左移动的次数, 进而 $S_n = \eta_n - (n - \eta_n) = 2\eta_n - n$.

故 $E(S_n)=E(2\eta_n-n)=2E(\eta_n)-n=2np-n$.

B5. 抛一枚均匀的硬币, 直到正、反两面都出现后停止试验, 求试验的平均次数.

解: 设试验次数为 X. 此时 X 可能取值为 $2,3,4,\cdots$

则 $\{X=n\}$ 表示前 $n-1$ 次结果均相同, 第 n 次和前 $n-1$ 次结果相反. 又因为有前正后反和前反后正两种情况, 则 $P\{X=n\}=2\cdot\left(\frac{1}{2}\right)^{n-1}\cdot\frac{1}{2}=\frac{1}{2^{n-1}}$.

从而平均次数即数学期望 $E(X)=\sum\limits_{n=2}^{+\infty}nP\{X=n\}=\sum\limits_{n=2}^{+\infty}n\left(\frac{1}{2}\right)^{n-1}$.

记 $S_n=\sum\limits_{i=2}^{n}i\left(\frac{1}{2}\right)^{i-1}$, $n\geqslant 2$. 则 $n=2$ 时 $S_2=2\times\frac{1}{2}=1$.

$n\geqslant 3$ 时, 有

$$\begin{aligned}\frac{1}{2}S_n=\sum_{i=2}^{n}i\left(\frac{1}{2}\right)^{i}&=\sum_{i=3}^{n+1}(i-1)\left(\frac{1}{2}\right)^{i-1}=\sum_{i=3}^{n+1}i\left(\frac{1}{2}\right)^{i-1}-\sum_{i=3}^{n+1}\left(\frac{1}{2}\right)^{i-1}\\&=(n+1)\cdot\left(\frac{1}{2}\right)^{n}+S_n-2\times\frac{1}{2}-\frac{\left(\frac{1}{2}\right)^2\left[1-\left(\frac{1}{2}\right)^{n-1}\right]}{1-\frac{1}{2}}\\&=\frac{n+1}{2^n}+S_n-1-\left(\frac{1}{2}-\frac{1}{2^n}\right)\\&=S_n+\frac{n+2}{2^n}-\frac{3}{2}.\end{aligned}$$

从而 $\frac{1}{2}S_n=\frac{3}{2}-\frac{n+2}{2^n}$, 则 $S_n=3-\frac{n+2}{2^{n-1}}$.

又 $n=2$ 时 $3-\frac{2+2}{2^1}=1=S_2$, 因此有 $S_n=3-\frac{n+2}{2^{n-1}}$, $n\geqslant 2$, $n\in\mathbb{N}$.

此时

$$\lim_{n\to+\infty}S_n=3-\lim_{n\to+\infty}\frac{n+2}{2^{n-1}}=3-0=3.$$

从而平均次数

$$E(X)=\lim_{n\to+\infty}S_n=3.$$

B6. 设二维随机变量 (X,Y) 的联合密度函数为

$$f(x,y)=\begin{cases}\dfrac{2}{x}e^{-2x}, & 0<x<+\infty, 0<y<x,\\ 0, & \text{其他},\end{cases}$$

(1) 求 $E(X)$; (2) 求 $E(3X-1)$; (3) 求 $E(XY)$.

解: (1) 由题

$$
\begin{aligned}
E(X) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xf(x,y)\mathrm{d}x\mathrm{d}y = \int_0^{+\infty}\mathrm{d}x\int_0^x x\cdot\frac{2}{x}e^{-2x}\mathrm{d}y = \int_0^{+\infty} 2e^{-2x}y\Big|_0^x \mathrm{d}x \\
&= \int_0^{+\infty} 2xe^{-2x}\mathrm{d}x = -\int_0^{+\infty} x\mathrm{d}e^{-2x} = -xe^{-2x}\Big|_0^{+\infty} + \int_0^{+\infty} e^{-2x}\mathrm{d}x \\
&= 0 - 0 - \frac{1}{2}e^{-2x}\Big|_0^{+\infty} = -\frac{1}{2}(0-1) = \frac{1}{2}.
\end{aligned}
$$

(2) $E(3X-1) = 3E(X) - 1 = \frac{1}{2}$.

(3) 由题, 利用 (1) 的结论有

$$
\begin{aligned}
E(XY) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xyf(x,y)\mathrm{d}x\mathrm{d}y = \int_0^{+\infty}\mathrm{d}x\int_0^x 2y\mathrm{e}^{-2x}\mathrm{d}y = \int_0^{+\infty} y^2\mathrm{e}^{-2x}\Big|_0^x \mathrm{d}x \\
&= \int_0^{+\infty} x^2e^{-2x}\mathrm{d}x = -\frac{1}{2}\int_0^{+\infty} x^2\mathrm{d}e^{-2x} = -\frac{1}{2}\left(x^2e^{-2x}\Big|_0^{+\infty} - \int_0^{+\infty} e^{-2x}\mathrm{d}x^2\right) \\
&= -\frac{1}{2}(0-0) + \frac{1}{2}\int_0^{+\infty} 2xe^{-2x}\mathrm{d}x = 0 + \frac{1}{2}\cdot\frac{1}{2} = \frac{1}{4}.
\end{aligned}
$$

B7. 已知一根长度为 1 的棍子上有个标志点 Q, 现随机地将此棍子折成两段.

(1) 已知点 Q 距离棍子某一端点的距离为 q, 求包含点 Q 的那一段棍子的平均长度(若截点刚好是点 Q, 则认为点 Q 包含在较短的一段中);

(2) 当点 Q 位于棍子何处时, 包含点 Q 的棍子平均长度达到最大?

解: (1) 设截断点到某一端点的距离为 X, 则 $X \sim U(0,1)$. 则其密度函数 $f_X(x) = \begin{cases} 1, & 0 < x < 1, \\ 0, & \text{其他}. \end{cases}$

记包含 Q 的棍子长度为 Y. 从而 $Y = g(X) = \begin{cases} X, & X > q, \\ \min\{X, 1-X\}, & X = q, \\ 1-X, & X < q. \end{cases}$

因此有

$$
\begin{aligned}
E(Y) &= \int_{-\infty}^{+\infty} g(x)f_X(x)\mathrm{d}x = \int_0^q (1-x)\mathrm{d}x + \int_q^1 x\mathrm{d}x = \left(x - \frac{1}{2}x^2\right)^2\Big|_0^q + \frac{1}{2}x^2\Big|_q^1 \\
&= q - \frac{1}{2}q^2 + \frac{1}{2} - \frac{1}{2}q^2 = \frac{1}{2} + q - q^2.
\end{aligned}
$$

(2) 平均长度最大时, $E(Y)$ 最大. 此时由于 $E(Y)$ 是二次函数, 则 $q = -\frac{1}{-2} = \frac{1}{2}$ 时 $E(Y)$ 取最大值. 因此 Q 点在棍子中点时平均长度最大.

B8. 甲、乙两人约定上午 8:00 ~ 9:00 在某地见面, 两人均在该时段随机到达, 且到达时间相互独立, 求两

人中先到的人需要等待的平均时间.

解: 设 8:00 后过 X 分钟甲到, 过 Y 分钟乙到. 则 $X \sim U(0,60)$, $Y \sim U(0,60)$.

X, Y 的密度函数分别为

$$f_X(x) = \begin{cases} \frac{1}{60}, & 0 < x < 60, \\ 0, & \text{其他}. \end{cases} \qquad f_Y(y) = \begin{cases} \frac{1}{60}, & 0 < y < 60, \\ 0, & \text{其他}. \end{cases}$$

由于 X, Y 相互独立, 则联合密度函数

$$f(x,y) = f_X(x) \cdot f_Y(y) = \begin{cases} \frac{1}{3600}, & 0 < x < 60, 0 < y < 60, \\ 0, & \text{其他}. \end{cases}$$

因为先到的人等待的时间 $Z = g(X,Y) = |X - Y| = \begin{cases} Y - X, & 0 < X < Y < 60, \\ X - Y, & 0 < Y < X < 60, \\ 0, & \text{其他}. \end{cases}$

因此有

$$\begin{aligned} E(Z) = E(g(X,Y)) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x,y) f(x,y) \mathrm{d}x \mathrm{d}y \\ &= \int_0^{60} \mathrm{d}x \int_0^x \frac{x-y}{3600} \mathrm{d}y + \int_0^{60} \mathrm{d}x \int_x^{60} \frac{y-x}{3600} \mathrm{d}y \\ &= \frac{1}{3600} \int_0^{60} \left(xy - \frac{1}{2} y^2 \right) \Big|_0^x \mathrm{d}x + \frac{1}{3600} \int_0^{60} \left(\frac{1}{2} y^2 - xy \right) \Big|_x^{60} \mathrm{d}x \\ &= \frac{1}{3600} \int_0^{60} \frac{1}{2} x^2 \mathrm{d}x + \frac{1}{3600} \int_0^{60} \left(\frac{1}{2} x^2 - 60x + 1800 \right) \Big|_0^x \mathrm{d}x \\ &= \frac{1}{3600} \cdot \frac{1}{6} x^3 \Big|_0^{60} + \frac{1}{3600} \cdot \left(\frac{1}{6} x^3 - 30x^2 + 1800x \right) \Big|_0^{60} \\ &= \frac{60^3}{3600 \times 6} + \frac{1}{3600} \left(\frac{1}{6} \times 60^3 - 30 \times 60^2 + 1800 \times 60 \right) \\ &= 10 + \frac{36000 - 108000 + 108000}{3600} = 10 + 10 = 20. \end{aligned}$$

即先到的人需要等待的平均时间为 20 分钟.

B9. 为诊断 500 人是否有人患有某种疾病, 抽血化验, 可用两种方法: (1) 每个人化验一次; (2) 分成 k 人一组 $\left(\text{共 } \frac{500}{k} \text{ 组, 假设 } \frac{500}{k} \text{ 为正整数}, k > 1 \right)$, 将每组 k 人的血液集中起来一起检验, 若化验结果为阴性, 则说明组内的每人都是阴性, 就无须分别化验; 若检验结果为阳性, 则说明这 k 人中至少有一人患病, 那么就对该组内的 k 人再单独化验. 如果此病的得病率为 20%, 问哪种方法的平均检验次数相对少一些?

解: 此时考虑方法(2)的平均次数.

设该方法中, 需要进行再次检测的组数为 X.

在某一组内, 该组 k 人全部为阴性的概率为 0.8^k, 因此该小组阳性需要再次检测的概率为 $1-0.8^k$.

从而对 X 有 $X\sim B\left(\dfrac{500}{k},1-0.8^k\right)$. 因此 $E(X)=\dfrac{500}{k}(1-0.8^k)$.

从而再次检测的平均组数为 $E(X)$, 则再次检测的平均次数即为平均人数 $kE(X)=500(1-0.8^k)$.

因此方法(2)的平均次数为 $\dfrac{500}{k}+500(1-0.8^k)$.

当 $\dfrac{500}{k}+500(1-0.8^k)=500$ 时, $\dfrac{1}{k}=0.8^k$, 即 $k0.8^k=1$.

对 $f(k)=k0.8^k-1$, 因为 $f'(k)=0.8^k+k\ln 0.8\cdot 0.8^{k-1}=0.8^{k-1}(0.8+k\ln 0.8)$. 注意到 $\ln 0.8<0$.

$f'(k)=0$ 时 $k=\dfrac{0.8}{\ln 5-\ln 4}=k_0\approx 4.481$. $k<k_0$ 时 $f'(k)>0$; $k>k_0$ 时 $f'(k)<0$.

又因为 $k>1$, 则 $k=2$ 时 $f(2)=2\times 0.8^2-1=0.28>0$. 即 $2<k<k_0$ 时 $f(k)>0$ 恒成立.

当 $k>k_0$ 时 $f(k)$ 单调递减, 且 $f(10)=10\times 0.8^{10}-1\approx 0.074>0$, $f(11)=11\times 0.8^{11}-1\approx -0.055<0$.

因此 $f(k)$ 在 $(10,11)$ 上存在零点, 且当 $k\leqslant 10$ 时, $k0.8^k>1$.

此时 $\dfrac{1}{k}<0.8^k$, 有 $\dfrac{500}{k}+500(1-0.8^k)<500$, 故方法(2)平均次数少.

又因为 $\dfrac{500}{k}$ 为正整数, 且 $k>1$. 因此 k 只能取 $2,4,5,10,20,25,50,100,125,250,500$.

因此当 $k=2,4,5,10$ 时采用方法(2)检验次数更少;

$k=20,25,50,100,125,250,500$ 时采用方法(1)检验次数更少.

B10. 某设备无故障运行的时间 T (以 h 计) 服从数学期望为 $\dfrac{1}{\lambda}(\lambda>0)$ 的指数分布. 若设备在一天 8 h 的工作时间内发生故障就自动停止运行待次日检修, 否则就运行 8 h 后停止, 求该设备每天运行的平均时间.

解: 由题 $T\sim E(\lambda)$, 则密度函数 $f_T(t)=\begin{cases}\lambda e^{-\lambda t}, & t>0,\\ 0, & t\leqslant 0.\end{cases}$

设每天运行的平均时间为 X, 由题可得 $X=\min\{T,8\}$, 且 $X>0$. 则

$$\begin{aligned}E(X)&=\int_{-\infty}^{+\infty}\min(t,8)f_T(t)\mathrm{d}t=\int_0^8\lambda te^{-\lambda t}\mathrm{d}t+\int_8^{+\infty}8\lambda e^{-\lambda t}\mathrm{d}t\\&=-\int_0^8 t\mathrm{d}e^{-\lambda t}-8e^{-\lambda t}\Big|_8^{+\infty}=-te^{-\lambda t}\Big|_0^8+\int_0^8 e^{-\lambda t}\mathrm{d}t-(0-8e^{-8\lambda})\\&=-8e^{-8\lambda}-0-\frac{1}{\lambda}e^{-\lambda t}\Big|_0^8+8e^{-8\lambda}=-\frac{1}{\lambda}(e^{-8\lambda}-1)=\frac{1-e^{-8\lambda}}{\lambda}.\end{aligned}$$

B11. 某电子监视器的圆形屏幕半径为 r $(r>0)$, 若目标出现的位置点 A 服从均匀分布. 以圆形屏幕的圆心为原点, 设点 A 的平面直角坐标为 (x,y).

(1) 求 $E(X)$ 与 $E(Y)$;

(2) 求点 A 与屏幕中心位置 $(0,0)$ 的平均距离.

解: (1) 由均匀分布, 因为圆面积为 πr^2, 则 X, Y 的联合密度函数 $f(x,y)=\begin{cases}\dfrac{1}{\pi r^2}, & x^2+y^2<r^2,\\ 0, & \text{其他}.\end{cases}$

考虑圆内, 当 $-r<x<r$ 时, $-\sqrt{r^2-x^2}<y<\sqrt{r^2-x^2}$, 此时 X 的边际密度函数

$$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}y=\int_{-\sqrt{r^2-x^2}}^{\sqrt{r^2-x^2}}\frac{1}{\pi r^2}\mathrm{d}y=\frac{2\sqrt{r^2-x^2}}{\pi r^2}.$$

且其他情况下 $f_X(x)=0$, 从而

$$E(X)=\int_{-\infty}^{+\infty}xf_X(x)\mathrm{d}x=\int_{-r}^{r}\frac{2x\sqrt{r^2-x^2}}{\pi r^2}\mathrm{d}x=0.$$

这里利用了奇函数的性质.[$2x\sqrt{r^2-x^2}$ 为奇函数]

同理可得 $E(Y)=0$.

(2) 点 A 到屏幕中心的距离 $Z=g(X,Y)=\sqrt{X^2+Y^2}$. 从而平均距离

$$\begin{aligned}E(Z)=E(g(X,Y))&=\iint\limits_{x^2+y^2<r^2}g(x,y)f(x,y)\mathrm{d}x\mathrm{d}y=\frac{1}{\pi r^2}\iint\limits_{x^2+y^2<r^2}\sqrt{x^2+y^2}\mathrm{d}x\mathrm{d}y\\&=\frac{1}{\pi r^2}\int_0^{2\pi}\mathrm{d}\theta\int_0^r\rho\cdot\rho\mathrm{d}\rho=\frac{2\pi}{\pi r^2}\cdot\frac{1}{3}\rho^3\Big|_0^r=\frac{2}{r^2}\cdot\frac{r^3}{3}=\frac{2r}{3}.\end{aligned}$$

B12. 一个袋子中有 15 个均匀的球, 其中 a 个是白球, 其他的是黑球. 无放回地随机抽取 n 次 (每次取一球), 记取到的白球数为 ξ_n. 已知 $E(\xi_2)=\dfrac{4}{3}$.

(1) 求 a; (2) 求 $E(\xi_9)$.

解: (1) 因为 $E(\xi_2)>1$, 则白球一定不止一个. 则抽两次时, 白球可能个数为 $0,1,2$.

从而

$$P\{\xi_2=0\}=\frac{\mathrm{C}_{15-a}^2}{\mathrm{C}_{15}^2}=\frac{(15-a)(14-a)}{210},\quad P\{\xi_2=1\}=\frac{\mathrm{C}_a^1\mathrm{C}_{15-a}^1}{\mathrm{C}_{15}^2}=\frac{2a(15-a)}{210},$$

$$P\{\xi_2=2\}=\frac{\mathrm{C}_a^2}{\mathrm{C}_{15}^2}=\frac{a(a-1)}{210}.$$

因此有

$$E(\xi_2)=0\cdot P\{\xi_2=0\}+1\cdot P\{\xi_2=1\}+2\cdot P\{\xi_2=2\}=\frac{2a(15-a)+2a(a-1)}{210}=\frac{14a}{105}=\frac{2a}{15}=\frac{4}{3}.$$

因此 $a=10$.

(2) 由于袋中有 10 个白球, 5 个黑球, 则抽 9 次时白球最少为 $9-5=4$ 个, 最多为 9 个.

而此时 $P\{\xi_9=k\}=\dfrac{C_{10}^k C_5^{9-k}}{C_{15}^9}$, $k=4,5,6,7,8,9$.

由于 $C_{n-1}^{k-1}=\dfrac{(n-1)!}{[(n-1)-(k-1)]!(k-1)!}=\dfrac{(n-1)!}{(n-k)!(k-1)!}=\dfrac{k\cdot n!}{n\cdot(n-k)!k!}=\dfrac{k}{n}\cdot C_n^k$.

也即 $kC_n^k=nC_{n-1}^{k-1}$.

从而

$$
\begin{aligned}
E(\xi_9)&=\sum_{k=4}^{9}k\cdot\frac{C_{10}^k C_5^{9-k}}{C_{15}^9}=\frac{1}{C_{15}^9}\sum_{k=4}^{9}10C_9^{k-1}C_5^{9-k}\\
&=\frac{10}{C_{15}^9}\left(C_9^3C_5^5+C_9^4C_5^4+C_9^5C_5^3+C_9^6+C_5^2+C_9^7C_5^1+C_9^8C_5^0\right)\\
&=\frac{10}{C_{15}^9}C_{14}^8 \qquad \text{[后面相当于是9个白球5个红球中抽取8个球的所有可能性的和]}\\
&=10\cdot\frac{9!6!}{15!}\cdot\frac{14!}{8!6!}=\frac{10\times 9}{15}=6.
\end{aligned}
$$

B13. 从 $1\sim 100$ 这 100 个数中无放回地取 10 个数, 计算这 10 个数的和的数学期望.

解: 设抽到的数为 X_i, $i=1,2,\cdots,10$.

此时由于这 10 个数同时抽取, 每个数都等可能出现, 此时 $E(X_i)=\dfrac{1}{100}\displaystyle\sum_{i=1}^{100}i=\dfrac{1}{100}\times\dfrac{100\times 101}{2}=50.5$.

从而 10 个数之和的数学期望

$$
E(X)=\sum_{i=1}^{10}E(X_i)=10\times 50.5=505.
$$

B14. 有 n 张各不相同的卡片, 采用有放回抽样, 每次取一张, 共取 n 次, 则有些卡片会被取到, 甚至被取到很多次, 但有些卡片可能不曾被取到. 设这 n 张卡片中被取到的共有 X 张, 计算 $E(X)$, 并计算当 $n\to+\infty$ 时, $E\left(\dfrac{X}{n}\right)$ 的极限.

解: 对第 i 张卡片引入随机变量 Y_i, 当其被抽到时 $Y_i=1$, 否则 $Y_i=0$, $i=1,2,\cdots,n$.

$Y_i=0$ 时相当于每次都从其他 $n-1$ 张中抽取, $P\{Y_i=0\}=\left(\dfrac{n-1}{n}\right)^n$.

则 $E(Y_i)=P\{Y_i=1\}=1-\left(\dfrac{n-1}{n}\right)^n=\dfrac{n^n-(n-1)^n}{n^n}$.

因此有

$$
E(X)=\sum_{i=1}^{n}E(Y_i)=n\cdot\frac{n^n-(n-1)^n}{n^n}=\frac{n^n-(n-1)^n}{n^{n-1}}.
$$

此时

$$
\lim_{n\to+\infty}E\left(\frac{X}{n}\right)=\lim_{n\to+\infty}\frac{n^n-(n-1)^n}{n^n}=1-\lim_{n\to+\infty}\left(1-\frac{1}{n}\right)^n=1-\lim_{n\to+\infty}\left(1-\frac{1}{n}\right)^{-n\cdot(-1)}=1-e^{-1}.
$$

B15. 在区间 $(0,1)$ 中随机地取 $n(n\geqslant 2)$ 个点, 求相距最远的两个点的距离的数学期望.

解: 由题, 设第 i 个点对应值为 X_i, $i=1,2,\cdots,n$. 则 $X_i\sim U(0,1)$.

相距最远的两个点的距离即值最大的点和值最小的点的距离.

即对 $M=\max\{X_1,X_2,\cdots,X_n\}$, $N=\min\{X_1,X_2,\cdots,X_n\}$, 所求 $Y=M-N$.

因为 X 的分布函数 $F_X(x)=\begin{cases}0, & x\leqslant 0,\\ x, & 0<x<1,\\ 1, & x\geqslant 1.\end{cases}$

从而 M 的分布函数 $F_M(x)=(F_X(x))^n=\begin{cases}0, & x\leqslant 0,\\ x^n, & 0<x<1,\\ 1, & x\geqslant 1.\end{cases}$

N 的分布函数 $F_N(x)=1-[1-F_X(x)]^n=\begin{cases}0, & x\leqslant 0,\\ 1-(1-x)^n, & 0<x<1,\\ 1, & x\geqslant 1.\end{cases}$

M 的密度函数 $f_M(x)=F_M'(x)=\begin{cases}nx^{n-1}, & 0<x<1,\\ 0, & \text{其他}.\end{cases}$

N 的密度函数 $f_N(x)=F_N'(x)=\begin{cases}n(1-x)^{n-1}, & 0<x<1,\\ 0, & \text{其他}.\end{cases}$

因此

$$E(M)=\int_0^1 x\cdot nx^{n-1}\mathrm{d}x=\frac{n}{n+1}x^{n+1}\bigg|_0^1=\frac{n}{n+1}.$$

$$\begin{aligned}E(N)&=\int_0^1 x\cdot n(1-x)^{n-1}\mathrm{d}x \xlongequal{t=1-x} \int_1^0 n(1-t)t^{n-1}\mathrm{d}(1-t)=\int_0^1 n(t^{n-1}-t^n)\mathrm{d}t\\&=\left(t^n-\frac{n}{n+1}t^{n+1}\right)\bigg|_0^1=1-\frac{n}{n+1}=\frac{1}{n+1}.\end{aligned}$$

因此 $E(Y)=E(M)-E(N)=\dfrac{n-1}{n+1}$.

B16. 设进入大型购物中心的顾客有可能去其中的一家冷饮店购买冷饮, 购买的概率为 p $(0<p<1)$. 若在一天的营业时间内进入该购物中心的顾客数 X 服从参数为 $\lambda(\lambda>0)$ 的泊松分布, 求这一天去该冷饮店购买冷饮的顾客数 Y 的分布及数学期望.

解: 当 $X=n$ 时, $Y\sim B(n,p)$. 则对 $k=1,2,\cdots,n$ 有 $P\{Y=k|X=n\}=\mathrm{C}_n^k p^k(1-p)^{n-k}$.

则条件期望 $E(Y|X=n)=np$.

又由泊松分布, $P\{X=n\}=\dfrac{e^{-\lambda}\lambda^n}{n!}$, 由全期望公式

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

B17. 设二维随机变量 (X,Y) 的联合密度函数为

$$f(x,y)=\begin{cases}\dfrac{3}{4}x, & 0<x<y<2,\\ 0, & 其他,\end{cases}$$

(1) 求 $E(Y|X=x)$;

(2) 用两种方法计算 $E(Y)$.

解: (1) 由于 $0<x<2$ 时 $f_X(x)=\displaystyle\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}y=\int_x^2\frac{3}{4}x\mathrm{d}y=\frac{3}{4}x(2-x)$,

从而当 $0<x<2$ 时, $f_{Y|X}(y|x)=\dfrac{f(x,y)}{f_X(x)}=\dfrac{1}{2-x}$.

因此

$$E(Y|X=x)=\int_{-\infty}^{+\infty}yf_{Y|X}(y|x)\mathrm{d}y=\int_x^2\frac{y}{2-x}\mathrm{d}y=\left.\frac{y^2}{2(2-x)}\right|_x^2=\frac{2^2-x^2}{2(2-x)}=\frac{x+2}{2}.$$

(2) 【方法一】直接利用二重积分计算. 此时有

$$\begin{aligned}E(Y)&=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}yf(x,y)\mathrm{d}x\mathrm{d}y=\int_0^2\mathrm{d}y\int_0^y\frac{3}{4}xy\mathrm{d}x=\frac{3}{4}\int_0^2\left.\frac{1}{2}yx^2\right|_0^y\mathrm{d}y\\&=\frac{3}{4}\int_0^2\frac{1}{2}y^3\mathrm{d}y=\left.\frac{3}{32}y^4\right|_0^2=\frac{3}{32}\times2^4=\frac{3}{2}.\end{aligned}$$

【方法二】利用全期望公式 $E(Y)=E[E(Y|X)]$. 有

$$\begin{aligned}E(Y)&=E[E(Y|X)]=\int_{-\infty}^{+\infty}E(Y|X=x)f_X(x)\mathrm{d}x=\int_0^2\frac{x+2}{2}\cdot\frac{3}{4}x(2-x)\mathrm{d}x\\&=\frac{3}{8}\int_0^2x(4-x^2)\mathrm{d}x=\left.\frac{3}{8}\left(2x^2-\frac{1}{4}x^4\right)\right|_0^2=\frac{3}{8}(2^3-2^2)=\frac{3}{2}.\end{aligned}$$

B18. (接第 B12 题) 当 $n=2$ 时, 求 $\mathrm{Var}(\xi_2)$.

解: 由 B12, $a=10$, 代入则有 $P\{\xi_2=0\}=\dfrac{(15-10)\times(14-10)}{210}=\dfrac{2}{21}$;

$P\{\xi_2=1\}=\dfrac{2\times10\times(15-10)}{210}=\dfrac{10}{21}$; $P\{\xi_2=2\}=\dfrac{10\times9}{210}=\dfrac{9}{21}=\dfrac{3}{7}$.

因此

$$\begin{aligned}\mathrm{Var}(\xi_2) &= \frac{2}{21}\times\left(0-\frac{4}{3}\right)^2+\frac{10}{21}\times\left(1-\frac{4}{3}\right)^2+\frac{9}{21}\times\left(2-\frac{4}{3}\right)^2\\ &= \frac{32}{189}+\frac{10}{189}+\frac{36}{189}=\frac{78}{189}=\frac{26}{63}.\end{aligned}$$

B19. 已知随机变量 X 服从 Γ 分布, 密度函数为

$$f(x)=\frac{\lambda^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\lambda x}, x>0, \alpha>0, \lambda>0.$$

其中 α 称为“形状参数”, λ 称为“尺度参数”, 求 $E(X^k)(k\geqslant 1)$ 和 $\mathrm{Var}(X)$.

解: 由于

$$\begin{aligned}E(X^k) &= \int_0^{+\infty}x^kf(x)\mathrm{d}x=\int_0^{+\infty}\frac{\lambda^\alpha}{\Gamma(\alpha)}x^{k+\alpha-1}e^{-\lambda x}\mathrm{d}x=\frac{\lambda^\alpha}{\lambda^{k+\alpha}\Gamma(\alpha)}\int_0^{+\infty}(\lambda x)^{k+\alpha-1}e^{-\lambda x}\mathrm{d}(\lambda x)\\ &= \frac{1}{\lambda^k\Gamma(\alpha)}\cdot\Gamma(k+\alpha)(k\geqslant 1).\end{aligned}$$

所以 $E(X)=\dfrac{\Gamma(1+\alpha)}{\lambda\Gamma(\alpha)}=\dfrac{\alpha}{\lambda}$, $E(X^2)=\dfrac{\Gamma(2+\alpha)}{\lambda^2\Gamma(\alpha)}=\dfrac{\alpha(\alpha+1)}{\lambda^2}$.

因此 $\mathrm{Var}(X)=E(X^2)-(E(X))^2=\dfrac{\alpha(\alpha+1)}{\lambda^2}-\dfrac{\alpha^2}{\lambda^2}=\dfrac{\alpha}{\lambda^2}$.

B20. 设随机变量 X 服从拉普拉斯分布, 密度函数为

$$f(x)=\frac{1}{2}e^{-|x|}, -\infty<x<+\infty,$$

计算 X 与 $|X|$ 的方差.

解: 由于 $\displaystyle\int_0^{+\infty}xe^{-x}\mathrm{d}x=\Gamma(2)=1!=1$, $\displaystyle\int_0^{+\infty}x^2e^{-x}\mathrm{d}x=\Gamma(3)=2!=2$.

且 $f(x)$ 为偶函数, 则 $xf(x)$ 为奇函数, $|x|f(x)$ 与 $x^2f(x)$ 为偶函数. 从而

$E(X)=\displaystyle\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x=0$;

$E(X^2)=\displaystyle\int_{-\infty}^{+\infty}x^2f(x)\mathrm{d}x=2\int_0^{+\infty}x^2f(x)\mathrm{d}x=\int_0^{+\infty}x^2e^{-x}\mathrm{d}x=2$.

则 $\mathrm{Var}(X)=E(X^2)-(E(X))^2=2-0=2$.

$E(|X|)=\displaystyle\int_{-\infty}^{+\infty}|x|f(x)\mathrm{d}x=2\int_0^{+\infty}|x|f(x)\mathrm{d}x=\int_0^{+\infty}xe^{-x}\mathrm{d}x=1$;

$E(|X|^2)=E(X^2)=2$.

则 $\mathrm{Var}(|X|)=E(|X|^2)-(E(|X|))^2=2-1=1$.

B21. 机器处于不同状态时制造产品的质量有所差异. 如果机器运作正常, 那么产品的正品率为 98%; 如果机器老化, 那么产品的正品率为 90%; 如果机器处于需要维修的状态, 那么产品的正品率为 74%. 机器正常

运作的概率为 0.7, 老化的概率为 0.2, 需要维修的概率为 0.1. 现随机抽取了 100 件产品(假设生产这些产品的机器的状态相互独立).

(1) 求产品中非正品数的数学期望与方差;

(2) 在已知这些产品都是正常机器制造出来的条件下, 求正品数的数学期望和方差.

解: (1) 因为正品率为 $0.7 \times 98\% + 0.2 \times 90\% + 0.1 \times 74\% = 0.94$,

设非正品数为 X, 则非正品率为 $1 - 0.94 = 0.06$, 则 $X \sim B(100, 0.06)$.

从而数学期望 $E(X) = 100 \times 0.06 = 6$; 方差 $\mathrm{Var}(X) = 100 \times 0.06 \times (1 - 0.06) = 5.64$.

(2) 设正品数为 Y. 因为这些产品都是正常机器制造出来的, 则$Y \sim B(100, 0.98)$.

从而数学期望$E(Y) = 100 \times 0.98 = 98$; 方差 $\mathrm{Var}(Y) = 100 \times 0.98 \times (1 - 0.98) = 1.96$.

B22. 设随机变量 X 与 Y 独立同分布, 且都服从参数为 $\frac{1}{2}$ 的 0-1 分布.

(1) 求 $P\{X + Y \geqslant 1\}$;

(2) 计算 $E(X \cdot (-1)^Y)$ 及 $\mathrm{Var}(X \cdot (-1)^Y)$.

解: (1) 由于 $P\{X = 0\} = P\{X = 1\} = P\{Y = 0\} = P\{Y = 1\} = \frac{1}{2}$, 则 $X + Y$ 只能取 $0, 1, 2$.

又因为 X 与 Y 独立, 则

$$\begin{aligned} P\{X + Y \geqslant 1\} &= 1 - P\{X + Y = 0\} = 1 - P\{X = Y = 0\} \\ &= 1 - P\{X = 0\} \cdot P\{Y = 0\} = 1 - \frac{1}{2} \cdot \frac{1}{2} = \frac{3}{4}. \end{aligned}$$

(2) 令 $Z = X \cdot (-1)^Y$, 则 Z 可能取 $-1, 0, 1$.

$Z = 0$ 时, $X = 0$, Y 任取, 有 $P\{Z = 0\} = P\{X = 0\} = \frac{1}{2}$.

$Z = 1$ 时, $X = 1$, $Y = 0$. 有 $P\{Z = 1\} = P\{X = 1, Y = 0\} = P\{X = 1\} \cdot P\{Y = 0\} = \frac{1}{4}$.

$Z = -1$ 时, $X = 1$, $Y = 1$.有 $P\{Z = -1\} = P\{X = 1, Y = 1\} = P\{X = 1\} \cdot P\{Y = 1\} = \frac{1}{4}$.

因此 $E(X \cdot (-1)^Y) = E(Z) = 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{4} + (-1) \cdot \frac{1}{4} = 0$.

$\mathrm{Var}(X \cdot (-1)^Y) = \mathrm{Var}(Z) = (0 - 0)^2 \times \frac{1}{2} + (1 - 0)^2 \times \frac{1}{4} + (-1 - 0)^2 \times \frac{1}{4} = \frac{1}{2}$.

B23. 设系统 L 由两个相互独立的子系统 L_1 和 L_2 组成, L_1 和 L_2 的寿命 X 与 Y 分别服从数学期望为 $\frac{1}{2}, \frac{1}{4}$ 的指数分布, 就下列三种连接方式写出系统 L 的寿命 Z 的数学期望和变异系数:

(1) L_1 和 L_2 串联;

(2) L_1 和 L_2 并联;

(3) L_2 为 L_1 的备用.

解: 由于指数分布的期望是参数的倒数, 则有 $X \sim E(2)$, $Y \sim E(4)$.

因此 X 的分布函数 $F_X(x)=\begin{cases}1-\mathrm{e}^{-2x}, & x>0,\\ 0, & x\leqslant 0.\end{cases}$ Y 的分布函数 $F_Y(y)=\begin{cases}1-\mathrm{e}^{-4y}, & y>0,\\ 0, & y\leqslant 0.\end{cases}$

(1) L_1 与 L_2 串联时, $Z=\min(X,Y)$.

则 Z 的分布函数 $F_Z(t)=1-[1-F_X(t)][1-F_Y(t)]$.

因为 $t>0$ 时 $F_Z(t)=1-(\mathrm{e}^{-2t})(\mathrm{e}^{-4t})=1-\mathrm{e}^{-6t}$; $t\leqslant 0$ 时 $F_Z(t)=1-(1-0)\times(1-0)=0$.

则 $F_Z(t)=\begin{cases}1-\mathrm{e}^{-6t}, & t>0,\\ 0, & t\leq 0.\end{cases}$ 从而有 $Z\sim E(6)$.

因此 $E(Z)=\dfrac{1}{6}$, $\mathrm{Var}(Z)=\dfrac{1}{36}$. 则变异系数 $Cv(Z)=\dfrac{\sqrt{\mathrm{Var}(Z)}}{E(Z)}=\dfrac{\frac{1}{6}}{\frac{1}{6}}=1$.

(2) L_1 与 L_2 并联时, $Z=\max(X,Y)$.

因此 Z 的分布函数 $F_Z(t)=F_X(t)\cdot F_Y(t)=\begin{cases}(1-e^{-2t})(1-e^{-4t})=e^{-6t}-e^{-4t}-e^{-2t}+1, & t>0,\\ 0, & t\leqslant 0.\end{cases}$

从而密度函数 $f_Z(t)=F_Z'(t)=\begin{cases}-6e^{-6t}+4e^{-4t}+2e^{-2t}, & t>0,\\ 0, & t\leqslant 0.\end{cases}$

因为

$$\begin{aligned}\int_0^{+\infty} t\cdot 6e^{-6t}\mathrm{d}t &= -\int_0^{+\infty} t\mathrm{d}e^{-6t}=-te^{-6t}\Big|_0^{+\infty}+\int_0^{+\infty} e^{-6t}\mathrm{d}t\\ &=0-\frac{1}{6}e^{-6t}\Big|_0^{+\infty}=-\left(0-\frac{1}{6}\right)=\frac{1}{6}.\end{aligned}$$

$$\begin{aligned}\int_0^{+\infty} t^2\cdot 6e^{-6t}\mathrm{d}t &= -\int_0^{+\infty} t^2\mathrm{d}e^{-6t}=-t^2e^{-6t}\Big|_0^{+\infty}+\int_0^{+\infty} e^{-6t}\mathrm{d}t^2\\ &=0+2\int_0^{+\infty} te^{-6t}\mathrm{d}t=2\times\frac{1}{6}\times\frac{1}{6}=\frac{1}{18}.\end{aligned}$$

类似地有

$$\int_0^{+\infty} t\cdot 4e^{-4t}\mathrm{d}t=\frac{1}{4},\qquad \int_0^{+\infty} t\cdot 2e^{-2t}\mathrm{d}t=\frac{1}{2}.$$

$$\int_0^{+\infty} t^2\cdot 4e^{-4t}\mathrm{d}t=2\times\frac{1}{4}\times\frac{1}{4}=\frac{1}{8},\qquad \int_0^{+\infty} t^2\cdot 2e^{-2t}\mathrm{d}t=2\times\frac{1}{2}\times\frac{1}{2}=\frac{1}{2}.$$

从而

$$E(Z)=\int_0^{+\infty} t(-6e^{-6t}+4e^{-4t}+2e^{-2t})\mathrm{d}t=-\frac{1}{6}+\frac{1}{4}+\frac{1}{2}=\frac{7}{12}.$$

$$E(Z^2)=\int_0^{+\infty} t^2(-6e^{-6t}+4e^{-4t}+2e^{-2t})\mathrm{d}t=-\frac{1}{18}+\frac{1}{8}+\frac{1}{2}=\frac{-4+9+36}{72}=\frac{41}{72}.$$

从而 $\mathrm{Var}(Z)=E(Z^2)-(E(Z))^2=\dfrac{41}{72}-\dfrac{49}{144}=\dfrac{33}{144}$.

则变异系数 $Cv(Z)=\dfrac{\sqrt{\mathrm{Var}(X)}}{E(X)}=\dfrac{\sqrt{\dfrac{33}{144}}}{\dfrac{7}{12}}=\dfrac{\sqrt{33}}{7}$.

(3) L_2 为 L_1 的备用时, $Z=X+Y$.

因为 $f_X(x)=\begin{cases}2e^{-2x}, & x>0,\\ 0, & x\leqslant 0.\end{cases}$ $\quad f_Y(y)=\begin{cases}4e^{-4y}, & y>0,\\ 0, & y\leqslant 0.\end{cases}$

由于 X 与 Y 独立, 则联合密度函数 $f(x,y)=f_X(x)\cdot f_Y(y)=\begin{cases}8e^{-2x-4y}, & x>0,y>0,\\ 0, & \text{其他}.\end{cases}$

此时 $f(x,t-x)=e^{-2x-4(t-x)}=e^{2x-4t}$, $0<x<t$. 其他情况下 $f(x,t-x)=0$.

从而 $t>0$ 时

$$f_Z(t)=\int_{-\infty}^{+\infty}f(x,t-x)\mathrm{d}x=\int_0^t 8e^{2x-4t}\mathrm{d}x=4e^{2x-4t}\Big|_0^t=4(e^{-2t}-e^{-4t}).$$

则 $f_Z(t)=\begin{cases}4(\mathrm{e}^{-2t}-\mathrm{e}^{-4t}), & t>0,\\ 0, & t\leqslant 0.\end{cases}$ 因此由(2)中计算结果可得

$$E(Z)=\int_{-\infty}^{+\infty}tf_Z(t)\mathrm{d}t=\int_0^{+\infty}4te^{-2t}\mathrm{d}t-\int_0^{+\infty}4te^{-4t}\mathrm{d}t=2\times\frac{1}{2}-\frac{1}{4}=\frac{3}{4}.$$

$$E(Z^2)=\int_{-\infty}^{+\infty}t^2f_Z(t)\mathrm{d}t=\int_0^{+\infty}4t^2e^{-2t}\mathrm{d}t-\int_0^{+\infty}4t^2e^{-4t}\mathrm{d}t=2\times\frac{1}{2}-\frac{1}{8}=\frac{7}{8}.$$

从而 $\mathrm{Var}(Z)=E(Z^2)-(E(Z))^2=\dfrac{7}{8}-\dfrac{9}{16}=\dfrac{5}{16}$.

变异系数 $Cv(Z)=\dfrac{\sqrt{\mathrm{Var}(Z)}}{E(Z)}=\dfrac{\sqrt{\dfrac{5}{16}}}{\dfrac{3}{4}}=\dfrac{\sqrt{5}}{3}$.

B24. (接第 B20 题) (1) 求 X 与 $|X|$ 的相关系数, 并判断两者是否相关;

(2) 判断 X 与 $|X|$ 是否相互独立.

解: (1) 由题, 因为 $f(x)$ 为偶函数, 则 $x|x|f(x)$ 为奇函数.

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

因此 $\mathrm{Cov}(X,|X|)=E(X|X|)-E(X)E(|X|)=0$.

则 X 与 $|X|$ 不相关.

(2) 任取 $0<b<a$, 此时

$$P\{X<a\}=\int_{-\infty}^{a}\frac{1}{2}e^{-|x|}\mathrm{d}x=\frac{1}{2}\left(\int_{-\infty}^{0}e^{x}\mathrm{d}x+\int_0^a e^{-x}\mathrm{d}x\right)=\frac{1}{2}\left(e^{x}\Big|_{-\infty}^{0}-e^{-x}\Big|_0^a\right)$$

$$= \frac{1}{2}[1 - 0 - (e^{-a} - 1)] = 1 - \frac{e^{-a}}{2}.$$

$$P\{|X| < b\} = P\{-b < |X| < b\} = \int_{-b}^{b} \frac{1}{2} e^{-|x|} \mathrm{d}x = 2 \cdot \frac{1}{2} \int_{0}^{b} e^{-x} \mathrm{d}x = -e^{-x}\Big|_0^b = 1 - e^{-b}.$$

$$P\{X < a, |X| < b\} = P\{X < a \text{且} -b < X < b\} = P\{-b < X < b\} = 1 - e^{-b}.$$

则此时有 $P\{X < a, |X| < b\} \neq P\{X < a\} \cdot P\{|X| < b\}$.

从而 X 与 $|X|$ 不相互独立.

B25. 设随机变量 (X, Y) 的联合密度函数为

$$f(x,y) = \begin{cases} \frac{1}{4}(1 + xy), & |x| < 1, |y| < 1, \\ 0, & \text{其他}, \end{cases}$$

(1) 计算 X 与 Y 的相关系数,并判断它们的独立性和相关性;

(2) 计算 X^2 与 Y^2 的相关系数, 并判断它们的独立性和相关性.

解: (1) $|x| < 1$ 时, $f_X(x) = \int_{-1}^{1} \frac{1}{4}(1 + xy)\mathrm{d}y = 2\int_0^1 \frac{1}{4}\mathrm{d}y = \frac{1}{2}$.

其他时, $f_X(x) = 0$.

$|y| < 1$ 时, $f_Y(y) = \int_{-1}^{1} \frac{1}{4}(1 + xy)\mathrm{d}x = 2\int_0^1 \frac{1}{4}\mathrm{d}x = \frac{1}{2}$.

其他时, $f_Y(y) = 0$.

此时 $E(X) = \int_{-1}^{1} \frac{1}{2} x\mathrm{d}x = 0$; $E(X^2) = \int_{-1}^{1} \frac{1}{2}x^2\mathrm{d}x = \int_0^1 x^2\mathrm{d}x = \frac{1}{3}$.

$E(Y) = \int_{-1}^{1} \frac{1}{2} y\mathrm{d}y = 0$; $E(Y^2) = \int_{-1}^{1} \frac{1}{2}y^2\mathrm{d}y = \int_0^1 y^2\mathrm{d}y = \frac{1}{3}$.

$$\begin{aligned} E(XY) &= \int_{-1}^{1}\int_{-1}^{1} \frac{1}{4}xy(1 + xy)\mathrm{d}x\mathrm{d}y = \frac{1}{4}\int_{-1}^{1} \left(\frac{1}{2}yx^2 + \frac{1}{3}y^2x^3\right)\Big|_{-1}^{1} \mathrm{d}y \\ &= \frac{1}{4}\int_{-1}^{1} \left[\left(\frac{1}{2}y + \frac{1}{3}y^2\right) - \left(\frac{1}{2}y - \frac{1}{3}y^2\right)\right] \mathrm{d}y = \frac{1}{4}\int_{-1}^{1} \frac{2}{3}y^2\mathrm{d}y \\ &= \frac{1}{3}\int_0^1 y^2\mathrm{d}y = \frac{1}{3} \cdot \frac{1}{3} = \frac{1}{9}. \end{aligned}$$

又因为 $\mathrm{Var}(X) = E(X^2) - (E(X))^2 = \frac{1}{3}$; $\mathrm{Var}(Y) = E(Y^2) - (E(Y))^2 = \frac{1}{3}$.

则 $\rho_{XY} = \dfrac{E(XY) - E(X)E(Y)}{\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}} = \dfrac{\frac{1}{9} - 0}{\sqrt{\frac{1}{3}} \cdot \sqrt{\frac{1}{3}}} = \frac{1}{3}$.

所以 X 和 Y 相关. 因此 X 和 Y 不独立.

(2)

$$\begin{aligned}E(X^2Y^2)&=\int_{-1}^{1}\int_{-1}^{1}\frac{1}{4}x^2y^2(1+xy)\mathrm{d}x\mathrm{d}y=\frac{1}{4}\int_{-1}^{1}\left(\frac{1}{3}y^2x^3+\frac{1}{4}y^3x^4\right)\Big|_{-1}^{1}\mathrm{d}y\\&=\frac{1}{4}\int_{-1}^{1}\left[\left(\frac{1}{3}y^2+\frac{1}{4}y^3\right)-\left(-\frac{1}{3}y^2+\frac{1}{4}y^3\right)\right]\mathrm{d}y=\frac{1}{4}\int_{-1}^{1}\frac{2}{3}y^2\mathrm{d}y\\&=\frac{1}{3}\int_{0}^{1}y^2\mathrm{d}y=\frac{1}{3}\cdot\frac{1}{3}=\frac{1}{9}.\end{aligned}$$

因此 $\mathrm{Cov}(X^2,Y^2)=E(X^2Y^2)-E(X^2)E(Y^2)=\dfrac{1}{9}-\dfrac{1}{3}\cdot\dfrac{1}{3}=0$,

则 $\rho_{X^2Y^2}=0$, X^2 与 Y^2 不相关.

再考虑独立性.

设 X^2 的边际分布函数为 $F_{X^2}(x)=P\{X^2\leqslant x\}$.

当 $x\leqslant 0$ 时, $F_{X^2}(x)=P\{X^2\leqslant x\}=0$;

当 $0<x<1$ 时, $F_{X^2}(x)=P\{X^2\leqslant x\}=P\{-\sqrt{x}\leqslant X\leqslant\sqrt{x}\}=\displaystyle\int_{-\sqrt{x}}^{\sqrt{x}}\frac{1}{2}\mathrm{d}t=\sqrt{x}$.

当 $x\geqslant 1$ 时, $F_{X^2}(x)=P\{X^2\leqslant x\}=1$.

因此 $F_{X^2}(x)=\begin{cases}0, & x\leqslant 0,\\ \sqrt{x}, & 0<x<1,\\ 1, & x\geqslant 1.\end{cases}$ 类似地, Y^2 的边际分布函数 $F_{Y^2}(y)=\begin{cases}0, & y\leqslant 0,\\ \sqrt{y}, & 0<y<1,\\ 1, & y\geqslant 1.\end{cases}$

设 (X^2,Y^2) 的联合分布函数 $F(x,y)=P\{X^2\leqslant x,Y^2\leqslant y\}$.

当 $0<x<1$, $0<y<1$ 时,

$$\begin{aligned}F(x,y)&=P\{X^2\leqslant x,Y^2\leqslant y\}=P\{-\sqrt{x}\leqslant X\leqslant\sqrt{x},-\sqrt{y}\leqslant Y\leqslant\sqrt{y}\}\\&=\int_{-\sqrt{y}}^{\sqrt{y}}\int_{-\sqrt{x}}^{\sqrt{x}}\frac{1}{4}(1+uv)\mathrm{d}u\mathrm{d}v=\int_{-\sqrt{y}}^{\sqrt{y}}\left(\frac{1}{2}\int_{0}^{\sqrt{x}}1\mathrm{d}u\right)\mathrm{d}v\\&=\frac{1}{2}\int_{-\sqrt{y}}^{\sqrt{y}}\sqrt{x}\mathrm{d}v=\sqrt{x}\int_{0}^{\sqrt{y}}1\mathrm{d}v=\sqrt{xy}.\end{aligned}$$

当 $0<x<1$, $y\geqslant 1$ 时, 则 $Y^2\leqslant y$ 即等价于 $|Y|\leqslant 1$. 从而有

$$\begin{aligned}F(x,y)&=P\{X^2\leqslant x,Y^2\leqslant y\}=P\{-\sqrt{x}\leqslant X\leqslant\sqrt{x},-1\leqslant Y\leqslant 1\}\\&=\int_{-1}^{1}\int_{-\sqrt{x}}^{\sqrt{x}}\frac{1}{4}(1+uv)\mathrm{d}u\mathrm{d}v=\int_{-1}^{1}\left(\frac{1}{2}\int_{0}^{\sqrt{x}}1\mathrm{d}u\right)\mathrm{d}v\\&=\frac{1}{2}\int_{-1}^{1}\sqrt{x}\mathrm{d}v=\sqrt{x}.\end{aligned}$$

当 $x\geqslant 1$, $0<y<1$ 时, 则 $X^2\leqslant x$ 即等价于 $|X|\leqslant 1$. 从而有

$$F(x,y)=P\{X^2\leqslant x,Y^2\leqslant y\}=P\{-1\leqslant X\leqslant 1,-\sqrt{y}\leqslant Y\leqslant\sqrt{y}\}$$

$$
\begin{aligned}
&= \int_{-\sqrt{y}}^{\sqrt{y}} \int_{-1}^{1} \frac{1}{4}(1+uv)\mathrm{d}u\mathrm{d}v = \int_{-\sqrt{y}}^{\sqrt{y}} \left(\frac{1}{2}\int_{0}^{1} 1\mathrm{d}u\right)\mathrm{d}v \\
&= \frac{1}{2}\int_{-\sqrt{y}}^{\sqrt{y}} 1\mathrm{d}v = \sqrt{y}.
\end{aligned}
$$

当 $x \geqslant 1, y \geqslant 1$ 时, $F(x,y) = P\{X^2 \leqslant x, Y^2 \leqslant y\} = 1$;

其他情况时, $F(x,y) = 0$.

因此 $F(x,y) = F_{X^2}(x) \cdot F_{Y^2}(y)$, 从而 X^2 和 Y^2 相互独立.

B26. 独立地抛一枚均匀的骰子 n 次 $(n \geqslant 2)$. 记 X, Y 分别为试验中"1 点朝上"以及"6 点朝上"出现的次数, 求 X 与 Y 的相关系数, 并判断两者的相关关系.

解: 由题, 有 $X \sim B\left(n, \frac{1}{6}\right)$, $Y \sim B\left(n, \frac{1}{6}\right)$.

因此有

$$
E(X) = E(Y) = \sum_{i=0}^{n} i\mathrm{C}_n^i \left(\frac{1}{6}\right)^i \left(\frac{5}{6}\right)^{n-i} = \frac{n}{6},
$$

$$
\mathrm{Var}(X) = \mathrm{Var}(Y) = \frac{1}{6}n \cdot \frac{5}{6} = \frac{5n}{36}.
$$

设 n 次中 1 点朝上次数为 i, 6 点朝上次数为 j. 此时有 $0 \leqslant i \leqslant n$, $0 \leqslant j \leqslant n-i$.则有

$$
P\{X=i, Y=j\} = \mathrm{C}_n^i \left(\frac{1}{6}\right)^i \left(\frac{5}{6}\right)^{n-i} \cdot \mathrm{C}_{n-i}^j \left(\frac{1}{5}\right)^j \left(\frac{4}{5}\right)^{n-i-j}
$$

由于对 $Z \sim B\left(n-i, \frac{1}{5}\right)$ 有

$$
E(Z) = \sum_{j=0}^{n-i} j \cdot \mathrm{C}_{n-i}^j \left(\frac{1}{5}\right)^j \left(\frac{4}{5}\right)^{n-i-j} = \frac{n-i}{5}.
$$

$P\{X=i\} = C_n^i$, 进而

$$
\begin{aligned}
E(XY) &= \sum_{i=0}^{n}\sum_{j=0}^{n-i} ij\mathrm{C}_n^i \left(\frac{1}{6}\right)^i \left(\frac{5}{6}\right)^{n-i} \mathrm{C}_{n-i}^j \left(\frac{1}{5}\right)^j \left(\frac{4}{5}\right)^{n-i-j} \\
&= \sum_{i=0}^{n} i\mathrm{C}_n^i \left(\frac{1}{6}\right)^i \left(\frac{5}{6}\right)^{n-i} \cdot \frac{n-i}{5} \\
&= \frac{1}{5}\sum_{i=0}^{n} i \cdot \frac{n!(n-i)}{(n-i)!i!} \left(\frac{1}{6}\right)^i \left(\frac{5}{6}\right)^{n-i} = \frac{1}{5}\sum_{i=0}^{n-1} i \cdot \frac{n!(n-i)}{(n-i)!i!} \left(\frac{1}{6}\right)^i \left(\frac{5}{6}\right)^{n-i} \quad [i=n\text{时取值为}0] \\
&= \frac{1}{5}\sum_{i=0}^{n-1} i \cdot \frac{n \cdot (n-1)!}{(n-i-1)!i!} \left(\frac{1}{6}\right)^i \left(\frac{5}{6}\right)^{n-i-1} \cdot \frac{5}{6} \\
&= \frac{n}{6}\sum_{i=0}^{n-1} i \cdot \frac{(n-1)!}{(n-1-i)!i!} \left(\frac{1}{6}\right)^i \left(\frac{5}{6}\right)^{n-1-i} \\
&= \frac{n}{6} \cdot \frac{n-1}{6} = \frac{n(n-1)}{36}.
\end{aligned}
$$

最后一行的求和式即为 $U \sim B\left(n-1,\dfrac{1}{6}\right)$ 的数学期望.

此时有

$$\mathrm{Cov}(X,Y)=E(XY)-E(X)E(Y)=\frac{n(n-1)}{36}-\frac{n}{6}\cdot\frac{n}{6}=\frac{n}{36}.$$

进而

$$\rho_{XY}=\frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}}=\frac{-\dfrac{n}{36}}{\sqrt{\dfrac{5n}{36}}\cdot\sqrt{\dfrac{5n}{36}}}=-\frac{1}{5}.$$

则 X 与 Y 负相关.

B27. 有一随机 △ABC, $\angle A$ 与 $\angle B$ 独立同分布, 设 A 的概率分布律如下:

A	$\frac{\pi}{3}$	$\frac{\pi}{4}$	$\frac{\pi}{6}$
p	λ	θ	$1-\lambda-\theta$

其中 $\lambda>0,\theta>0$, 且满足 $\lambda+\theta<1$. 已知 $E(\sin A)=E(\cos A)=\dfrac{\sqrt{3}+2\sqrt{2}+1}{8}$.

(1) 写出 (A,B) 的联合分布律;

(2) 求 $E(\sin C)$;

(3) 求 $\angle A$ 与 $\angle C$ 的相关系数, 并由此判断它们的相关性 (若相关, 说明是正相关还是负相关).

解: (1) 由于

$$E(\sin A)=\sin\frac{\pi}{3}\cdot\lambda+\sin\frac{\pi}{4}\cdot\theta+\sin\frac{\pi}{6}\cdot(1-\lambda-\theta)=\frac{1+(\sqrt{3}-1)\lambda+(\sqrt{2}-1)\theta}{2}.$$

$$E(\cos A)=\cos\frac{\pi}{3}\cdot\lambda+\cos\frac{\pi}{4}\cdot\theta+\cos\frac{\pi}{6}\cdot(1-\lambda-\theta)=\frac{\sqrt{3}+(1-\sqrt{3})\lambda+(\sqrt{2}-\sqrt{3})\theta}{2}.$$

又 $E(\sin A)=E(\cos A)=\dfrac{\sqrt{3}+2\sqrt{2}+1}{8}$, 二者相加则有 $\dfrac{1+\sqrt{3}}{2}+\dfrac{2\sqrt{2}-1-\sqrt{3}}{2}\theta=\dfrac{\sqrt{3}+2\sqrt{2}+1}{4}$.

即 $\dfrac{2\sqrt{2}-1-\sqrt{3}}{2}\theta=\dfrac{2\sqrt{2}-\sqrt{3}-1}{4}$, 则 $\theta=\dfrac{1}{2}$.

从而有 $\dfrac{4+4(\sqrt{3}-1)\lambda+2(\sqrt{2}-1)}{8}=\dfrac{\sqrt{3}+2\sqrt{2}+1}{8}$, 有 $4(\sqrt{3}-1)\lambda=\sqrt{3}+1-2$, 解得 $\lambda=\dfrac{1}{4}$.

即 $P\left\{A=\dfrac{\pi}{3}\right\}=\dfrac{1}{4}$; $P\left\{A=\dfrac{\pi}{4}\right\}=\dfrac{1}{2}$; $P\left\{A=\dfrac{\pi}{6}\right\}=\dfrac{1}{4}$.

又因为 A 与 B 独立同分布, 则 $P\{A=\alpha,B=\beta\}=P\{A=\alpha\}\cdot P\{B=\beta\}$.

且 B 的分布律为 $P\left\{B=\dfrac{\pi}{3}\right\}=\dfrac{1}{4}$; $P\left\{B=\dfrac{\pi}{4}\right\}=\dfrac{1}{2}$; $P\left\{B=\dfrac{\pi}{6}\right\}=\dfrac{1}{4}$.

因此计算得 (A,B) 的联合分布律为

A \ B	$\frac{\pi}{3}$	$\frac{\pi}{4}$	$\frac{\pi}{6}$
$\frac{\pi}{3}$	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{16}$
$\frac{\pi}{4}$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{8}$
$\frac{\pi}{6}$	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{16}$

(2) 由于 $\sin C=\sin(A+B)$, 此时只需考察 $A+B$ 的分布. 此时 $A+B$ 可以取 $\frac{\pi}{3},\frac{5\pi}{12},\frac{\pi}{2},\frac{7\pi}{12},\frac{2\pi}{3}$.

$$P\left\{A+B=\frac{\pi}{3}\right\}=P\left\{A=B=\frac{\pi}{6}\right\}=\frac{1}{16};$$

$$P\left\{A+B=\frac{2\pi}{3}\right\}=P\left\{A=B=\frac{\pi}{3}\right\}=\frac{1}{16};$$

$$P\left\{A+B=\frac{5\pi}{12}\right\}=P\left\{A=\frac{\pi}{4},B=\frac{\pi}{6}\right\}+P\left\{A=\frac{\pi}{6},B=\frac{\pi}{4}\right\}=\frac{1}{4};$$

$$P\left\{A+B=\frac{7\pi}{12}\right\}=P\left\{A=\frac{\pi}{4},B=\frac{\pi}{3}\right\}+P\left\{A=\frac{\pi}{3},B=\frac{\pi}{4}\right\}=\frac{1}{4};$$

$$P\left\{A+B=\frac{\pi}{2}\right\}=1-\frac{1}{4}-\frac{1}{4}-\frac{1}{16}-\frac{1}{16}=\frac{3}{8}.$$

又因为 $\sin\frac{5\pi}{12}=\sin\frac{7\pi}{12}=\sin\left(\frac{\pi}{4}+\frac{\pi}{6}\right)=\sin\frac{\pi}{4}\cos\frac{\pi}{6}+\sin\frac{\pi}{6}\cos\frac{\pi}{4}=\frac{\sqrt{6}+\sqrt{2}}{4}$, 从而

$$\begin{aligned}E(\sin C)=E(\sin(A+B))&=\sin\frac{\pi}{3}\cdot\frac{1}{16}+\sin\frac{5\pi}{12}\cdot\frac{1}{4}+\sin\frac{\pi}{2}\cdot\frac{3}{8}+\sin\frac{7\pi}{12}\cdot\frac{1}{4}+\sin\frac{2\pi}{3}\cdot\frac{1}{16}\\&=\frac{\sqrt{3}}{32}+\frac{\sqrt{6}+\sqrt{2}}{16}+\frac{3}{8}+\frac{\sqrt{6}+\sqrt{2}}{16}+\frac{\sqrt{3}}{32}\\&=\frac{6+2\sqrt{6}+\sqrt{3}+2\sqrt{2}}{16}.\end{aligned}$$

(3) 因为 $E(A)=E(B)=\frac{\pi}{3}\cdot\frac{1}{4}+\frac{\pi}{4}\cdot\frac{1}{2}+\frac{\pi}{6}\cdot\frac{1}{4}=\frac{\pi}{12}+\frac{\pi}{8}+\frac{\pi}{24}=\frac{\pi}{4}$.

$\operatorname{Var}(A)=\operatorname{Var}(B)=\frac{1}{4}\cdot\left(\frac{\pi}{3}-\frac{\pi}{4}\right)^2+\frac{1}{2}\cdot\left(\frac{\pi}{4}-\frac{\pi}{4}\right)^2+\frac{1}{4}\cdot\left(\frac{\pi}{6}-\frac{\pi}{4}\right)^2=\frac{\pi^2}{576}+\frac{\pi^2}{576}=\frac{\pi^2}{288}$.

又因为 $C=\pi-A-B$, 且 A 与 B 独立, 有

$$\operatorname{Var}(C)=\operatorname{Var}(\pi-A-B)=\operatorname{Var}(A+B)=\operatorname{Var}(A)+\operatorname{Var}(B)=\frac{\pi^2}{144}.$$

$$\operatorname{Cov}(A,C)=\operatorname{Cov}(A,\pi-A-B)=0-\operatorname{Cov}(A,A)-\operatorname{Cov}(A,B)=-\operatorname{Var}(A)-0=-\frac{1}{288}\pi^2.$$

因此相关系数

$$\rho_{AC}=\frac{\mathrm{Cov}(A,C)}{\sqrt{\mathrm{Var}(A)}\sqrt{\mathrm{Var}(C)}}=\frac{-\frac{1}{288}\pi^2}{\sqrt{\frac{\pi^2}{288}\cdot\frac{\pi^2}{144}}}=-\frac{\frac{1}{288}\pi^2}{\frac{\pi^2}{144\sqrt{2}}}=-\frac{\sqrt{2}}{2}.$$

从而 A 与 C 是负相关.

B28. 设随机变量 $X_1,X_2,\cdots,X_n$ 均服从标准正态分布并且相互独立. 记 $S_k=\sum\limits_{i=1}^{k}X_i$, $T_k=\sum\limits_{j=n_0+1}^{n_0+k}X_j$, 其中 $1\leqslant n_0<k<n_0+k\leqslant n$. 求 S_k 与 T_k 的相关系数.

解: 由题有 $E(X_i)=0$, $\mathrm{Var}(X_i)=1$, $i=1,2,\cdots,n$.

因此有 $E(S_k)=E\left(\sum\limits_{i=1}^{k}X_i\right)=\sum\limits_{i=1}^{k}E(X_i)=0$, $E(T_k)=E\left(\sum\limits_{j=n_0+1}^{n_0+k}X_j\right)=\sum\limits_{j=n_0+1}^{n_0+k}E(X_j)=0$.

又由相互独立知

$$\mathrm{Var}(S_k)=\mathrm{Var}\left(\sum_{i=1}^{k}X_i\right)=\sum_{i=1}^{k}\mathrm{Var}(X_i)=k\cdot 1=k.$$

$$\mathrm{Var}(T_k)=\mathrm{Var}\left(\sum_{j=n_0+1}^{n_0+k}X_j\right)=\sum_{j=n_0+1}^{n_0+k}\mathrm{Var}(X_j)=k\cdot 1=k.$$

此时因为 $E(X_iX_j)=E(X_i)E(X_j)=0$, $i\neq j$; $E(X_i^2)=\mathrm{Var}(X_i)+(E(X_i))^2=1$.

从而有

$$E(S_kT_k)=E\left(\sum_{i=1}^{k}X_i\sum_{j=n_0+1}^{n_0+k}X_j\right)=\sum_{i=1}^{k}\sum_{j=n_0+1}^{n_0+k}E(X_iX_j)=\sum_{i=n_0+1}^{k}E(X_i^2)=k-n_0.$$

故 $\mathrm{Cov}(S_k,T_k)=E(S_kT_k)-E(S_k)E(T_k)=k-n_0$. 从而

$$\rho_{S_kT_k}=\frac{\mathrm{Cov}(S_k,T_k)}{\sqrt{\mathrm{Var}(S_k)}\sqrt{\mathrm{Var}(T_k)}}=\frac{k-n_0}{k}.$$

B29. 设 $X\sim N(0,1)$, Y 的可能取值为 ±1, 且 $P\{Y=1\}=p(0<p<1)$. 设 X 与 Y 相互独立,记 $\xi=X\cdot Y$.

(1) 证明: $\xi\sim N(0,1)$;

(2) 计算 $\rho_{X\xi}$ 并判断 X 与 ξ 的相关性和独立性.

解: (1) 由全概率公式, 结合正态分布的性质有

$$\begin{aligned}P\{\xi\leqslant t\}&=P\{X\cdot Y\leqslant t\}=P\{X\cdot Y\leqslant t|Y=1\}P\{Y=1\}+P\{X\cdot Y\leqslant t|Y=-1\}P\{Y=-1\}\\&=pP\{X\leqslant t\}+(1-p)P\{-X\leqslant t\}=pP\{X\leqslant t\}+(1-p)P\{X\geqslant -t\}\\&=pP\{X\leqslant t\}+(1-p)P\{X\leqslant t\}=P\{X\leqslant t\}.\end{aligned}$$

因此 ξ 和 X 同分布, 则 $\xi\sim N(0,1)$.

(2) 由题, $E(Y)=1\cdot p+(-1)\cdot(1-p)=2p-1$, 且 $E(\xi)=E(X)=0$, $\mathrm{Var}(\xi)=\mathrm{Var}(X)=1$.

因为 X 与 Y 独立, 则 X^2 与 Y 独立. 又 $E(X^2)=\mathrm{Var}(X)+(E(X))^2=1$, 有

$$E(X\xi)=E(XXY)=E(X^2Y)=E(X^2)E(Y)=1\cdot(2p-1)=2p-1.$$

从而 $\mathrm{Cov}(X,\xi)=E(X\xi)-E(X)E(\xi)=2p-1$. 有

$$\rho_{X\xi}=\frac{\mathrm{Cov}(X,\xi)}{\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(\xi)}}=\frac{2p-1}{\sqrt{1}\times\sqrt{1}}=2p-1.$$

因此当 $0<p<0.5$ 时, X 和 ξ 负相关;

当 $p=0.5$ 时, X 和 ξ 不相关;

当 $0.5<p<1$ 时, X 和 ξ 正相关.

再考虑独立性.

由于 $P\{X\leqslant x,\xi\leqslant t\}=P\{X\leqslant x\}P\{\xi\leqslant t|X\leqslant x\}=P\{X\leqslant x\}P\{X\cdot Y\leqslant t|X\leqslant x\}\xlongequal{\text{由(1)}}P\{X\leqslant x\}P\{X\leqslant t|X\leqslant x\}$

当$t>x$时, $P\{X\leqslant x,\xi\leqslant t\}=P\{X\leqslant x\}\neq P\{X\leqslant x\}P\{\xi\leqslant t\}=P\{X\leqslant x\}P\{X\leqslant t\}$, 故$X$和$\xi$不独立.

B30. 有 n 包巧克力, 每包重量服从分布 $N(\mu,\sigma^2)$, 且各包重量相互独立, 求前 $k(1\leqslant k\leqslant n)$ 包重量与这 n 包总重量的相关系数.

解: 设第 i 包重量为 X_i, $i=1,2,\cdots,n$. 设前 k 包重量为 $Y=\sum\limits_{i=1}^{k}X_i$, n 包总重量为 $Z=\sum\limits_{j=1}^{n}X_j$.

此时 $X_i\sim N(\mu,\sigma^2)$. 则 $E(Y)=kE(X_1)=k\mu$, $E(Z)=nE(X_1)=n\mu$.

由各包重量独立, 则 $\mathrm{Var}(Y)=k\mathrm{Var}(X_1)=k\sigma^2$, $\mathrm{Var}(Z)=n\mathrm{Var}(X_1)=n\sigma^2$.

且 $i\neq j$ 时 $E(X_iX_j)=E(X_i)E(X_j)=\mu^2$, 从而

$$E(YZ)=E\left(\sum_{i=1}^{k}X_i\sum_{j=1}^{n}X_j\right)=kE(X_1^2)+(kn-k)E(X_1)E(X_1)$$
$$=k(\mu^2+\sigma^2)+(kn-k)\mu^2=kn\mu^2+k\sigma^2.$$

【k 项与 n 项多项式相乘共有 kn 项, 其中 k 项是相同项相乘】

从而相关系数

$$\mathrm{Cov}(Y,Z)=\frac{E(YZ)-E(Y)E(Z)}{\sqrt{\mathrm{Var}(Y)}\sqrt{\mathrm{Var}(Z)}}=\frac{kn\mu^2+k\sigma^2-k\mu\cdot n\mu}{\sqrt{k\sigma^2}\cdot\sqrt{n\sigma^2}}=\frac{k\sigma^2}{\sqrt{kn\sigma^2}}=\sqrt{\frac{k}{n}}.$$

B31. 设甲、乙两个盒子中都装有 2 个白球 3 个黑球. 先从甲盒中任取 1 个球放入乙盒, 再从乙盒中随机地取出一球, 用 X 与 Y 分别表示从甲、乙盒中取得的白球数.

(1) 求 (X,Y) 的联合分布律, 并判断 X 与 Y 是否相互独立;

(2) 求出 $\mathrm{Cov}(X,Y)$, 并由此判断 X 与 Y 的相关性.

解: (1) 此时 X 与 Y 可能取值均为 0,1.

先考虑取球情况, 甲中取到白球时, 乙中即为 3 白 3 黑; 甲中取到黑球时, 乙中即为 2 白 4 黑.

$\{X=0,Y=0\}$ 即甲黑乙黑, $P\{X=0,Y=0\}=\dfrac{3}{5}\times\dfrac{4}{6}=\dfrac{2}{5}$.

$\{X=0,Y=1\}$ 即甲黑乙白, $P\{X=0,Y=1\}=\dfrac{3}{5}\times\dfrac{2}{6}=\dfrac{1}{5}$.

$\{X=1,Y=0\}$ 即甲白乙黑, $P\{X=1,Y=0\}=\dfrac{2}{5}\times\dfrac{3}{6}=\dfrac{1}{5}$.

$\{X=1,Y=1\}$ 即甲白乙白, $P\{X=1,Y=1\}=\dfrac{2}{5}\times\dfrac{3}{6}=\dfrac{1}{5}$.

从而 (X,Y) 联合分布律为

X	Y	
	0	1
0	0.4	0.2
1	0.2	0.2

又因为 $P\{X=0\}=\dfrac{3}{5}=0.6$, $P\{Y=0\}=0.4+0.2=0.6$,

则 $P\{X=0,Y=0\}\neq P\{X=0\}\cdot P\{Y=0\}$, 从而 X 与 Y 不独立.

(2) 因为 $P\{X=1\}=P\{Y=1\}=0.4$, 则 $E(X)=E(Y)=0\cdot 0.6+1\cdot 0.4=0.4$.

又因为 $P\{XY=1\}=P\{X=1,Y=1\}=0.2$, $P\{XY=0\}=1-0.2=0.8$.

从而 $E(XY)=1\times 0.2+0\times 0.8=0.2$, 则 $\mathrm{Cov}(X,Y)=E(XY)-E(X)E(Y)=0.2-0.4\times 0.4=0.04$.

由于 $\mathrm{Cov}(X,Y)>0$, 则 X 与 Y 正相关.

B32. 设二维随机变量 (X,Y) 服从正态分布 $N(0,1;1,4;\rho)$. 令 $\xi=aX-bY$, $\eta=aY-bX$, 其中 a,b 为实数, $a\neq b$ 且 $ab\neq 0$.

(1) 当 $\rho=0$ 时, 分别写出 ξ 与 η 的分布 (要求写出参数) 及它们各自的标准化变量, 并计算 ξ 与 η 的相关系数;

(2) 当 $\rho=\dfrac{1}{2}$ 时, 计算 ξ 的变异系数;

(3) 当 $\rho=\dfrac{1}{2}$ 时, 计算 η 的中位数;

(4) 当 $\rho=-1$ 时, 判断 ξ 与 η 的独立性和相关性.

解: 由题, $X\sim N(0,1)$, $Y\sim N(1,4)$.【该条件适用于整个题目】

(1) 由 $\rho=0$ 则 X 与 Y 独立.

此时 $E(\xi)=E(aX-bY)=aE(X)-bE(Y)=-b$, $E(\eta)=E(aY-bX)=aE(Y)-bE(X)=a$.

$\mathrm{Var}(\xi)=\mathrm{Var}(aX-bY)=a^2\mathrm{Var}(X)+b^2\mathrm{Var}(Y)=a^2+4b^2$.

$\mathrm{Var}(\eta)=\mathrm{Var}(aY-bX)=a^2\mathrm{Var}(Y)+b^2\mathrm{Var}(X)=4a^2+b^2$.

且由于 ξ,η 均为正态分布的线性组合, 故 $\xi\sim N(-b,a^2+4b^2)$, $\eta\sim N(a,4a^2+b^2)$.

设二者的标准化变量分别为 ξ^* 与 η^*, 则 $\xi^*=\dfrac{\xi+b}{\sqrt{a^2+4b^2}}$, $\eta^*=\dfrac{\eta-a}{\sqrt{4a^2+b^2}}$.

因为 $E(X^2)=\mathrm{Var}(X)+(E(X))^2=1$, $E(Y^2)=\mathrm{Var}(Y)+(E(Y))^2=5$.

又因为 $E(XY)=E(X)E(Y)=0$, 则有

$$E(\xi\eta)=E[(aX-bY)(aY-bX)]=a^2E(XY)-abE(X^2)-abE(Y^2)+b^2E(XY)=-6ab.$$

故 $\mathrm{Cov}(\xi,\eta)=E(\xi\eta)-E(\xi)E(\eta)=-6ab-(-ab)=-5ab$. 因此

$$\rho_{\xi\eta}=\frac{\mathrm{Cov}(\xi,\eta)}{\sqrt{D(\xi)}\sqrt{D(\eta)}}=-\frac{5ab}{\sqrt{(a^2+4b^2)(4a^2+b^2)}}.$$

(2) 此时 $\mathrm{Cov}(X,Y)=\rho\cdot\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}=\dfrac{1}{2}\cdot\sqrt{1}\cdot\sqrt{4}=1$.

且此时 $E(\xi)=-b$,

$$\begin{aligned}\mathrm{Var}(\xi)&=\mathrm{Var}(aX-bY)=a^2\mathrm{Var}(X)+b^2\mathrm{Var}(Y)+2\ \mathrm{Cov}(aX,-bY)\\&=a^2+4b^2-2ab\ \mathrm{Cov}(X,Y)=a^2-2ab+4b^2.\end{aligned}$$

则变异系数 $Cv(\xi)=\dfrac{\sqrt{\mathrm{Var}(\xi)}}{E(\xi)}=-\dfrac{\sqrt{a^2-2ab+4b^2}}{b}$.

(3) 由于 η 服从正态分布, 则 η 的中位数即为其数学期望, 即 $\eta_{1/2}=a$.

(4) 由(1), $E(\xi)=-b$, $E(\eta)=a$. 且由于 η 与 ξ 均服从正态分布, 则二者若相关, 则必不独立.

又因为 $\mathrm{Cov}(X,Y)=\rho\cdot\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}=-1\cdot\sqrt{1}\cdot\sqrt{4}=-2$.

则 $E(XY)=\mathrm{Cov}(X,Y)+E(X)E(Y)=-2$. 又因为 $E(X^2)=1$, $E(Y^2)=5$, 则

$$E(\xi\eta)=E[(aX-bY)(aY-bX)]=a^2E(XY)-abE(X^2)-abE(Y^2)+b^2E(XY)=-6ab-2a^2-2b^2.$$

从而

$$\begin{aligned}\mathrm{Cov}(\xi,\eta)&=E(\xi\eta)-E(\xi)E(\eta)=-4ab-2a^2-2b^2-(-ab)\\&=-(2a^2+5ab+2b^2)=-(2a+b)(a+2b).\end{aligned}$$

因此 $b=-2a$ 或 $a=-2b$ 时, $\mathrm{Cov}(\xi,\eta)=0$, 则此时 ξ 与 η 不相关, 且相互独立.

又因为 $-\dfrac{(2a+b)(a+2b)}{b^2}=-\left(\dfrac{2a}{b}+1\right)\left(\dfrac{a}{b}+2\right)$, 则 $\mathrm{Cov}(\xi,\eta)>0$ 时, $\left(\dfrac{2a}{b}+1\right)\left(\dfrac{a}{b}+2\right)<0$.

此时即 $-2<\dfrac{a}{b}<-\dfrac{1}{2}$, 对应的 ξ 与 η 正相关, 且不独立;

其他情况时, ξ 与 η 负相关, 且不独立.

B33. 已知三维正态变量 $\boldsymbol{X}=(X_1,X_2,X_3)^{\mathrm{T}}\sim N(\boldsymbol{a},\boldsymbol{B})$, 其中 $\boldsymbol{a}=(0,0,1)^{\mathrm{T}}$, $\boldsymbol{B}=\begin{pmatrix}1&2&-1\\2&16&0\\-1&0&4\end{pmatrix}$.

(1) 写出 $\boldsymbol{X}$ 的每个分量的分布;

(2) 判别 X_1,X_2,X_3 的相关性与独立性;

(3) 若 $Y_1=X_1-X_2$, $Y_2=X_3-X_1$, 求 $\boldsymbol{Y}=(Y_1,Y_2)^{\mathrm{T}}$ 的分布;

解: (1) 三个分量记为 X_1, X_2, X_3. 由题可知, 每个分量都满足正态分布.

此时 $X_1\sim N(0,1)$, $X_2\sim N(0,16)$, $X_3\sim N(1,4)$.

(2) 由 $\boldsymbol{B}$ 可知, $\mathrm{Cov}(X_1,X_2)=2$, $\mathrm{Cov}(X_1,X_3)=-1$, $\mathrm{Cov}(X_2,X_3)=0$.

又因为 X_1, X_2, X_3 都满足正态分布, 从而相关即不独立, 不相关即独立.

因此 X_1 和 X_2 正相关且不独立; X_1 和 X_3 负相关且不独立; X_2 和 X_3 不相关且独立.

(3) 由线性组合可知, Y_1 和 Y_2 均服从正态分布.

此时

$$E(Y_1)=E(X_1-X_2)=E(X_1)-E(X_2)=0.$$

$$\mathrm{Var}(Y_1)=\mathrm{Var}(X_1-X_2)=\mathrm{Var}(X_1)+\mathrm{Var}(X_2)+2\mathrm{Cov}(X_1,-X_2)=1+16-2\times2=13.$$

$$E(Y_2)=E(X_3-X_1)=E(X_3)-E(X_1)=1.$$

$$\mathrm{Var}(Y_2)=\mathrm{Var}(X_3-X_1)=\mathrm{Var}(X_3)+\mathrm{Var}(X_1)+2\mathrm{Cov}(X_3,-X_1)=4+1-2\times(-1)=7.$$

$$\begin{aligned}E(Y_1Y_2)&=E[(X_1-X_2)(X_3-X_1)]=E(X_1X_3-X_1^2-X_2X_3+X_1X_2)\\&=E(X_1X_3)-E(X_1^2)-E(X_2X_3)+E(X_1X_2)\\&=\mathrm{Cov}(X_1,X_3)+E(X_1)E(X_3)-[\mathrm{Var}(X_1)+(E(X_1))^2]\\&\quad-[\mathrm{Cov}(X_2,X_3)+E(X_2)E(X_3)]+[\mathrm{Cov}(X_1,X_2)+E(X_1)E(X_2)]\\&=-1+0-(1+0)-(0+0)+(2+0)=0.\end{aligned}$$

从而 $\mathrm{Cov}(Y_1,Y_2)=E(Y_1Y_2)-E(Y_1)E(Y_2)=0-0=0$.

此时 $\boldsymbol{Y}=(Y_1,Y_2)^{\mathrm{T}}$ 的分布即为 $N(\boldsymbol{a}',\boldsymbol{B}')$, 其中 $\boldsymbol{a}'=(0,1)^{\mathrm{T}}$, $\boldsymbol{B}'=\begin{pmatrix}13&0\\0&7\end{pmatrix}$.

B34. 设有一煤矿一天的产煤量 X (以 10^4 t 计) 服从正态分布 $N(1.5, 0.1^2)$. 设每天产煤量相互独立,一个月按 30 天计. 求一个月总产煤量超过 46×10^4 t 的概率.

解: 设一个月总产煤量为 Y, 第 i 天产煤量 X_i, 则 $Y=\sum\limits_{i=1}^{30} X_i$.

由题, $X_i \sim N(1.5, 0.1^2)$, $1\leqslant i\leqslant 30$, 且 X_i 相互独立.

因此有 $Y\sim N(1.5\times 30, 0.1^2\times 30)$, 即$Y\sim N(45, 0.3)$. 从而 $\dfrac{Y-45}{\sqrt{0.3}}\sim N(0,1)$.

则

$$P\{Y\geqslant 46\}=1-P\{Y<46\}=1-P\left\{\frac{Y-45}{\sqrt{0.3}}<\frac{46-45}{\sqrt{0.3}}\right\}\approx 1-\Phi(1.83)=1-0.9664=0.0336.$$

【此题这里与课后答案有出入是因为计算时 $\sqrt{\dfrac{10}{3}}$ 在这里计算时近似两位得到 1.83, 并通过查书后附表 2 得到 $\Phi(1.83)$ 的值, 而没有直接计算 $\Phi\left(\sqrt{\dfrac{10}{3}}\right)$.】

B35. 某地区成年男子身高 X (单位: cm)服从正态分布 $N(170, 144)$, 从该地区独立抽选 4 人, 求这 4 人平均身高超过 176 cm 的概率.

解: 设抽取的 4 人的身高分别为 X_1, X_2, X_3, X_4.

则平均身高 $Y=\dfrac{1}{4}(X_1+X_2+X_3+X_4)$, 所求即为 $P\{Y\geqslant 176\}$.

又因为 X_1, X_2, X_3, X_4 相互独立, 且 $X_i\sim N(170, 144)$, $i=1,2,3,4$.

从而 $(X_1+X_2+X_3+X_4)\sim N(680, 576)$, 有 $Y\sim N(170, 36)$, 因此 $\dfrac{Y-170}{6}\sim N(0,1)$.

从而

$$\begin{aligned}P\{Y\geqslant 176\}&=1-P\{Y<176\}=1-P\left\{\frac{Y-170}{6}<\frac{176-170}{6}\right\}\\&=1-\Phi(1)=1-0.8413=0.1587.\end{aligned}$$

第 5 章　大数定律及中心极限定理

A1. 设 $X_1, X_2, \cdots, X_n, \cdots$ 相互独立, 均服从参数为 2 的指数分布. 问当 $n \to +\infty$ 时, $\frac{1}{n}\sum_{i=1}^{n} X_i^2$ 依概率收敛于何值?

解: 由于 $h(x) = x^2$ 连续, 且

$$
\begin{aligned}
E(h(X_1)) = E(X_1^2) &= \int_0^{+\infty} x^2 \cdot 2e^{-2x}\mathrm{d}x = -\int_0^{+\infty} x^2 \mathrm{d}e^{-2x} = -x^2e^{-2x}\Big|_0^{+\infty} + \int_0^{+\infty} e^{-2x}\mathrm{d}x^2 \\
&= 0 - 0 + 2\int_0^{+\infty} xe^{-2x}\mathrm{d}x = -\int_0^{+\infty} x\mathrm{d}e^{-2x} = -xe^{-2x}\Big|_0^{+\infty} + \int_0^{+\infty} e^{-2x}\mathrm{d}x \\
&= 0 - 0 - \frac{1}{2}e^{-2x}\Big|_0^{+\infty} = -\left(0 - \frac{1}{2}\right) = \frac{1}{2}.
\end{aligned}
$$

由辛钦大数定律, $\frac{1}{n}\sum_{i=1}^{n} X_i^2 \xrightarrow{P} E(X_1^2) = \frac{1}{2}$, $n \to +\infty$.

A2. 设随机变量序列 $\{X_n, n \geq 1\}$, 当 $n \to +\infty$ 时, X_n 依概率收敛于 3. 问当 $n \to +\infty$ 时, 下列随机变量序列依概率收敛于何值:

(1) X_n^2;　　　　(2) $2X_n - 3$.

解: 由依概率收敛的性质, 因为 $g(x) = x^2$, $h(x) = 2x - 3$ 在 $x = 3$ 处连续.

(1) $X_n^2 \xrightarrow{P} g(3) = 3^2 = 9$, $n \to +\infty$.

(2) $2X_n - 3 \xrightarrow{P} h(3) = 2 \times 3 - 3 = 3$, $n \to +\infty$.

A3. 设 X_1 与 X_2 相互独立, 均值都为 2, 方差都为 4, 用切比雪夫不等式求 $P\{|X_1 - X_2| \geqslant 4\}$ 的上界.

解: 由题

$$
E(X_1) = 2, \quad E(X_2) = 2; \qquad \mathrm{Var}(X_1) = 4, \quad \mathrm{Var}(X_2) = 4.
$$

因为 X_1 与 X_2 相互独立, 从而 $E(X_1 - X_2) = E(X_1) - E(X_2) = 2 - 2 = 0$,

$\mathrm{Var}(X_1 - X_2) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2) = 4 + 4 = 8$.

由切比雪夫不等式,

$P\{|X_1 - X_2| \geqslant 4\} = P\{|(X_1 - X_2) - 0| \geqslant 4\} \leqslant \frac{\mathrm{Var}(X_1 - X_2)}{4^2} = \frac{8}{16} = \frac{1}{2}$.

因此该概率的上限为 $\frac{1}{2}$.

A4. 设 $X_1, X_2, \cdots, X_{315}$ 独立同分布, X_1 的密度函数为

$$
f(x) = \begin{cases} \frac{2}{3}x, & 1 < x < 2, \\ 0, & \text{其他}. \end{cases}
$$

Y 表示 $\{X_i < 1.5\}$ $(i=1,2,\cdots,315)$ 出现的个数, 求 $P\{Y<140\}$ 的近似值.

解: 因为 $P\{X_1<1.5\} = \displaystyle\int_1^{1.5} \frac{2}{3}x\,\mathrm{d}x = \frac{1}{3}x^2\Big|_1^{1.5} = \frac{1}{3}\left(\frac{9}{4}-1\right) = \frac{5}{12}$.

因此对 Y 有 $Y \sim B\left(315, \dfrac{5}{12}\right)$.

由中心极限定理,

$$\frac{Y-315\times\frac{5}{12}}{\sqrt{315\times\frac{5}{12}\times\frac{7}{12}}} = \frac{Y-\frac{525}{4}}{\frac{35}{4}} = \frac{4Y-525}{35} \overset{\text{近似地}}{\sim} N(0,1).$$

所以

$$P\{Y<140\} = P\left\{\frac{4Y-525}{35} < \frac{560-525}{35}\right\} = P\left\{\frac{4Y-525}{35}<1\right\} \approx \Phi(1) = 0.8413.$$

A5. 设 $X_1, X_2, \cdots, X_{240}$ 独立同分布, $P\{X_1=-2\}=0.3$, $P\{X_1=0\}=0.4$, $P\{X_1=2\}=0.3$. 令 $Y=X_1+X_2+\cdots+X_{240}$, 求 $P\{|Y|>24\}$ 的近似值.

解: 由题 $E(X_1) = -2\times0.3+0\times0.4+2\times0.3=0$.

$\mathrm{Var}(X_1) = (-2-0)^2\times0.3+(0-0)^2\times0.4+(2-0)^2\times0.3 = 2.4$.

由中心极限定理,

$$\frac{Y-240\times0}{\sqrt{2.4}\times\sqrt{240}} = \frac{Y}{24} \overset{\text{近似地}}{\sim} N(0,1).$$

所以

$$\begin{aligned}P\{|Y|>24\} &= 1-P\{|Y|\leqslant 24\} = 1-(P\{Y\leqslant 24\}-P\{Y\leqslant -24\}) \\ &= 1-\left(P\left\{\frac{Y}{24}\leqslant 1\right\}-P\left\{\frac{Y}{24}\leqslant -1\right\}\right) \approx 1-(\Phi(1)-\Phi(-1)) \\ &= 1-(2\Phi(1)-1) = 2-2\Phi(1) = 2-1.6826 = 0.3174.\end{aligned}$$

B1. 某种类的昆虫每周产卵数为随机变量 X (以个计), 若已知其平均周产卵数为 36 个.

(1) 用马尔可夫不等式求一周内核昆虫产卵数不少于 50 个的概率的上界;

(2) 若又已知该昆虫每周产卵数的标准差为 2 个, 用切比雪夫不等式求一周内产卵数在 $(32,40)$ 内的概率的下界.

解: (1) 由题, $E(X)=36$, 故由马尔可夫不等式, $P\{X\geqslant 50\}\leqslant \dfrac{E(X)}{50}=\dfrac{36}{50}=0.72$.

即概率上界为 0.72.

(2) 由题, $\mathrm{Var}(X)=2^2=4$. 故由切比雪夫不等式,

$$P\{32<X<40\}=P\{-4<X-36<4\}=P\{|X-36|<4\}\geqslant 1-\frac{4}{4^2}=0.75.$$

即概率下界为 0.75.

B2. 一种遗传病的隔代发病率为 10%, 在得病家庭中选取 500 户进行研究, 试用切比雪夫不等式估计这 500 户中隔代发病的比例与发病率之差的绝对值小于 5% 的概率下界.

解: 用 X_i 表示第 i 户隔代发病的情况, 记 $X_i=1$ 为隔代发病, $X_i=0$ 为隔代不发病, 此时 $i=1,2,\cdots,500$.

由于发病率为 10%, 且每个家庭独立同分布, 则 $X_i\sim B(1,0.1)$. 此时有 $E(X_1)=0.1$, $\mathrm{Var}(X_1)=0.1\times 0.9=0.09$.

因此 $E\left(\dfrac{1}{500}\displaystyle\sum_{i=1}^{500}X_i\right)=0.1$, $\mathrm{Var}\left(\dfrac{1}{500}\displaystyle\sum_{i=1}^{500}X_i\right)=\dfrac{0.09\times 500}{500^2}=0.00018$.

从而由切比雪夫不等式

$$P\left\{\left|\frac{1}{500}\sum_{i=1}^{500}X_i-0.1\right|<0.05\right\}\geqslant 1-\frac{0.00018}{0.05^2}=1-0.072=0.928.$$

即概率下界为 0.928.

B3. 设随机变量 X_i 的密度函数为

$$f(x)=\begin{cases}\dfrac{i|x|^{i-1}}{2}, & |x|\leq 1,\\ 0, & \text{其他},\end{cases}\qquad i=1,2,\cdots,n.$$

且 $X_1,X_2,\cdots,X_n$ 相互独立. 令 $Y_n=X_1X_2\cdots X_n$, 用切比雪夫不等式求使得 $P\left\{|Y_n|\geqslant\dfrac{1}{2}\right\}\leqslant\dfrac{1}{9}$ 成立的最小的 n.

解: 由奇函数性质, $E(X_i)=\displaystyle\int_{-1}^{1}x\cdot\frac{i|x|^{i-1}}{2}\mathrm{d}x=0$, $i=1,2,\cdots,n$.

由独立则 $E(Y_n)=E(X_1X_2\cdots X_n)=E(X_1)E(X_2)\cdots E(X_n)=0$.

又因为

$$E(X_i^2)=\int_{-1}^{1}x^2\cdot\frac{i|x|^{i-1}}{2}\mathrm{d}x=2\int_0^1\frac{i}{2}x^{i+1}\mathrm{d}x=\frac{i}{(i+2)}x^{i+2}\Big|_0^1=\frac{i}{i+2}.$$

从而 $E(Y_n^2)=E(X_1^2)E(X_2^2)\cdots E(X_n^2)=\frac{1}{3}\cdot\frac{2}{4}\cdots\frac{n}{n+2}=\frac{2}{(n+1)(n+2)}$.

因此 $\mathrm{Var}(Y_n)=E(Y_n^2)+(E(Y_n))^2=\frac{2}{(n+1)(n+2)}$.

由切比雪夫不等式, $P\left\{|Y_n-0|\geqslant\frac{1}{2}\right\}\leqslant\frac{\mathrm{Var}(Y_n)}{0.5^2}=\frac{8}{(n+1)(n+2)}$.

因此有 $\frac{8}{(n+1)(n+2)}\leqslant\frac{1}{9}$, 则 $(n+1)(n+2)\geqslant 72$, 即 $n^2+3n-70=(n-7)(n+10)\geqslant 0$.

又因为 n 为正整数, 从而有 $n\geqslant 7$, 则最小 n 值为 7.

B4. 设随机变量序列 $\{X_n,n\geqslant 1\}$ 独立同分布, 都服从 $U(0,a)$, 其中 $a>0$. 令 $X_{(n)}=\max\limits_{1\leqslant i\leqslant n}X_i$, 证明: $X_{(n)}\xrightarrow{P}a,n\to+\infty$.

证明: 先确定 $X_{(n)}$ 的分布.

由于 $X_i\sim U(0,a)$, $1\leqslant i\leqslant n$, 故 X_i 的分布函数 $F_i(x)=\begin{cases}0, & x\leqslant 0,\\ \frac{x}{a}, & 0<x<a,\\ 1, & x\geqslant a.\end{cases}$

对 $X_{(n)}$, 其分布函数 $F_M(x)=\prod\limits_{i=1}^{n}F_i(x)$. 从而 $F_M(x)=\begin{cases}0, & x\leqslant 0,\\ \left(\frac{x}{a}\right)^n, & 0<x<a,\\ 1, & x\geqslant a.\end{cases}$

【方法一】利用依概率收敛的定义.

对任意 $\varepsilon>0$, 不妨设 $\varepsilon<a$. 因为 $X_{(n)}\leqslant a$ 恒成立, 有

$$\begin{aligned}\lim_{n\to+\infty}P\{|X_{(n)}-a|\geqslant\varepsilon\}&=\lim_{n\to+\infty}P\{a-X_{(n)}\geqslant\varepsilon\}=\lim_{n\to+\infty}P\{X_{(n)}\leqslant a-\varepsilon\}\\&=\lim_{n\to+\infty}\left(\frac{a-\varepsilon}{a}\right)^n=\lim_{n\to+\infty}\left(1-\frac{\varepsilon}{a}\right)^n=0.\end{aligned}$$

因此 $X_{(n)}\xrightarrow{P}a,n\to+\infty$.

【方法二】利用切比雪夫不等式.

由于 $0<x<a$ 时 $X_{(n)}$ 的密度函数 $f_M(x)=F_M'(x)=\frac{nx^{n-1}}{a^n}$, 其他时 $f_M(x)=0$.

则

$$E(X_{(n)})=\int_0^a x\cdot\frac{nx^{n-1}}{a^n}\mathrm{d}x=\frac{n}{a^n}\cdot\frac{x^{n+1}}{n+1}\Big|_0^a=\frac{n}{a^n}\cdot\frac{a^{n+1}}{n+1}=\frac{an}{n+1}.$$

$$E(X_{(n)}^2)=\int_0^a x^2\cdot\frac{nx^{n-1}}{a^n}\mathrm{d}x=\frac{n}{a^n}\cdot\frac{x^{n+2}}{n+2}\Big|_0^a=\frac{n}{a^n}\cdot\frac{a^{n+2}}{n+2}=\frac{a^2n}{n+2}.$$

$$\operatorname{Var}(X_{(n)})=E(X_{(n)}^2)-(E(X_{(n)}))^2=\frac{a^2n}{n+2}-\frac{a^2n^2}{(n+1)^2}=\frac{a^2[n(n+1)^2-n^2(n+2)]}{(n+1)^2(n+2)}=\frac{a^2n}{(n+1)^2(n+2)}.$$

因此 $\lim\limits_{n\to+\infty}E(X_{(n)})=a$, $\lim\limits_{n\to+\infty}\operatorname{Var}(X_{(n)})=0$.

由切比雪夫不等式, 对任意 $\varepsilon>0$, $P\left\{\left|X_{(n)}-E(X_{(n)})\right|\geqslant\varepsilon\right\}\leqslant\dfrac{\operatorname{Var}(X_{(n)})}{\varepsilon^2}$.

当 $n\to+\infty$ 时, 则有

$$\lim_{n\to+\infty}P\{|X_{(n)}-a|\geqslant\varepsilon\}\leqslant 0.$$

又因为 $P\{|X_{(n)}-a|\geqslant\varepsilon\}\geqslant 0$, 则 $\lim\limits_{n\to+\infty}P\{|X_{(n)}-a|\geqslant\varepsilon\}=0$.

因此 $X_{(n)}\xrightarrow{P}a, n\to+\infty$.

B5. 设随机变量序列 $\{X_i,i\geqslant 1\}$ 独立同分布, 数学期望与方差均存在, 证明:

$$\frac{2}{n(n+1)}\sum_{i=1}^n i\cdot X_i\xrightarrow{P}E(X_1),\qquad n\to+\infty.$$

解: 令 $Y=\sum\limits_{i=1}^n i\cdot X_i$, 因为 $\{X_i,i\geqslant 1\}$ 独立同分布, 设 $E(X_1)=\mu$, $\operatorname{Var}(X_1)=\sigma^2$, 则

$$E(Y)=E\left(\sum_{i=1}^n i\cdot X_i\right)=\sum_{i=1}^n E(i\cdot X_i)=\sum_{i=1}^n i\cdot E(X_i)=E(X_1)\sum_{i=1}^n i=\frac{n(n+1)}{2}\mu.$$

$$\begin{aligned}\operatorname{Var}(Y)=\operatorname{Var}\left(\sum_{i=1}^n i\cdot X_i\right)&=\sum_{i=1}^n\operatorname{Var}(iX_i)=\sum_{i=1}^n i^2\operatorname{Var}(X_i)\\&=\sigma^2\sum_{i=1}^n i^2=\frac{n(n+1)(2n+1)}{6}\sigma^2.\end{aligned}$$

此时对 $Z=\dfrac{2}{n(n+1)}Y$, $E(Z)=\mu$, $\operatorname{Var}(Z)=\dfrac{4}{n^2(n+1)^2}\cdot\dfrac{n(n+1)(2n+1)}{6}\sigma^2=\dfrac{2(2n+1)}{3n(n+1)}\sigma^2$.

从而由切比雪夫不等式, 对任意的 $\varepsilon>0$,

$$P\{|Z-E(Z)|\geqslant\varepsilon\}\leqslant\frac{\operatorname{Var}(Z)}{\varepsilon^2}=\frac{2(2n+1)}{3n(n+1)\varepsilon^2}\sigma^2<\frac{2(2n+2)}{3n(n+1)\varepsilon}\sigma^2=\frac{4\sigma^2}{3n\varepsilon^2}$$

因为 $n\to+\infty$ 时 $\dfrac{4\sigma^2}{3n\varepsilon^2}\to 0$. 则有

$$\lim_{n\to+\infty}P\left\{\left|\frac{2}{n(n+1)}\sum_{i=1}^n i\cdot X_i-E(X_1)\right|\geqslant\varepsilon\right\}=\lim_{n\to+\infty}P\{|Z-E(Z)|\geqslant\varepsilon\}=0.$$

从而 $\dfrac{2}{n(n+1)}\sum\limits_{i=1}^n i\cdot X_i\xrightarrow{P}E(X_1)$, $n\to+\infty$.

B6. 设 $\{X_i, i \geqslant 1\}$ 为独立同分布的正态随机变量序列, 若 $X_1 \sim N(\mu, \sigma^2)$, 其中 $\sigma > 0$. 问以下的随机变量序列当 $n \to +\infty$ 时依概率收敛吗? 若收敛, 请给出收敛的极限值, 否则请说明理由:

(1) $\frac{1}{n}\sum_{i=1}^{n} X_i^2$; (2) $\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2$;

(3) $\frac{X_1+X_2+\cdots+X_n}{X_1^2+X_2^2+\cdots+X_n^2}$; (4) $\frac{X_1+X_2+\cdots+X_n}{\sqrt{n\sum_{i=1}^{n}(X_i-\mu)^2}}$.

解: (1) 由于 $E(X_i^2) = \mathrm{Var}(X_i) + (E(X_i))^2 = \sigma^2 + \mu^2$, 且 $\{X_i^2, i \geqslant 1\}$ 独立同分布.

则由辛钦大数定律,

$$\frac{1}{n}\sum_{i=1}^{n} X_i^2 \xrightarrow{P} \sigma^2 + \mu^2, \quad n \to +\infty.$$

(2) 由于

$$E((X_i-\mu)^2) = E(X_i^2 - 2\mu X_i + \mu^2) = E(X_i^2) - 2\mu E(X_i) + \mu^2 = \sigma^2 + \mu^2 - 2\mu^2 + \mu^2 = \sigma^2.$$

由辛钦大数定律,

$$\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2 \xrightarrow{P} \sigma^2, \quad n \to +\infty.$$

(3) 由于 $E(X_i) = \mu$, 由辛钦大数定律 $\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{P} \mu$, $n \to +\infty$.

根据依概率收敛的性质, 对 $X_n \xrightarrow{P} a$, $Y_n \xrightarrow{P} b$, $g(x,y)$ 在 (a,b) 处连续, 则 $g(X_n, Y_n) \xrightarrow{P} g(a,b)$.

由(1), $\frac{1}{n}\sum_{i=1}^{n} X_i^2 \xrightarrow{P} \sigma^2 + \mu^2$, 而 $g(x,y) = \frac{x}{y}$ 在 $(\mu, \sigma^2+\mu^2)$ 处连续, 从而有

$$\frac{X_1+X_2+\cdots+X_n}{X_1^2+X_2^2+\cdots+X_n^2} = \frac{\frac{1}{n}\sum_{i=1}^{n} X_i}{\frac{1}{n}\sum_{i=1}^{n} X_i^2} \xrightarrow{P} \frac{\mu}{\sigma^2+\mu^2}, \quad n \to +\infty.$$

(4) 由上知 $\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2 \xrightarrow{P} \sigma^2$, 且 $\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{P} \mu$.

由依概率收敛的性质有

$$\frac{X_1+X_2+\cdots+X_n}{\sqrt{n\sum_{i=1}^{n}(X_i-\mu)^2}} = \frac{\frac{1}{n}\sum_{i=1}^{n} X_i}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2}} \xrightarrow{P} \frac{\mu}{\sigma}, \quad n \to +\infty.$$

B7. 设随机变量序列 $\{X_i, i \geqslant 1\}$ 独立同分布, 都服从期望为 $\frac{1}{\lambda}$ 的指数分布, 其中 $\lambda > 0$.

(1) 若对任意的 $\varepsilon>0$, 均有

$$\lim_{n\to+\infty} P\left\{\left|\frac{X_1^2+X_2^2+\cdots+X_n^2}{n}-a\right|<\varepsilon\right\}=1$$

成立, 求 a;

(2) 给出 $\frac{1}{50}\sum_{i=1}^{100} X_i$ 的近似分布;

(3) 求 $P\left\{\frac{1}{100}\sum_{i=1}^{100} X_i^2 \leqslant \frac{2}{\lambda^2}\right\}$ 的近似值.

解: (1) 由指数分布知, $E(X_1)=\frac{1}{\lambda}$, $\operatorname{Var}(X_1)=\frac{1}{\lambda^2}$.

而根据辛钦大数定律可得, $a=E(X_1^2)=\operatorname{Var}(X_1)+(E(X_1))^2=\frac{2}{\lambda^2}$.

(2) 由中心极限定理可得

$$\frac{\frac{1}{100}\sum\limits_{i=1}^{100} X_i-\frac{1}{\lambda}}{\frac{1}{\lambda}\cdot\frac{1}{\sqrt{100}}} \overset{\text{近似地}}{\sim} N(0,1).$$

从而有 $\frac{1}{100}\sum\limits_{i=1}^{100} X_i-\frac{1}{\lambda} \overset{\text{近似地}}{\sim} N\left(0,\frac{1}{100\lambda^2}\right)$, 则 $\sum\limits_{i=1}^{100} X_i \overset{\text{近似地}}{\sim} N\left(\frac{100}{\lambda},\frac{100}{\lambda^2}\right)$.

因此 $\frac{1}{50}\sum\limits_{i=1}^{100} X_i$ 的近似分布为 $N\left(\frac{2}{\lambda},\frac{1}{25\lambda^2}\right)$.

(3) 由于 $E(X_1^2)=\frac{2}{\lambda^2}$, 设 $\operatorname{Var}(X_1^2)=\sigma^2$, 由中心极限定理可得

$$\frac{\frac{1}{100}\sum\limits_{i=1}^{100} X_i^2-\frac{2}{\lambda^2}}{\frac{\sigma}{\sqrt{100}}} \overset{\text{近似地}}{\sim} N(0,1).$$

故 $P\left\{\frac{1}{100}\sum\limits_{i=1}^{100} X_i^2 \leqslant \frac{2}{\lambda^2}\right\}=P\left\{\frac{\frac{1}{100}\sum\limits_{i=1}^{100} X_i^2-\frac{2}{\lambda^2}}{\frac{\sigma}{\sqrt{100}}}\leqslant 0\right\}\approx\Phi(0)=0.5.$

B8. 抛掷一枚硬币 10000 次, 出现了 5325 次“正面”, 是否可以断言此硬币是不均匀的?

解: 记 n_A 为抛掷硬币 10000 次时出现“正面”的次数.

假设硬币是均匀的, 则 $n_A\sim B(10000,0.5)$, 由中心极限定理,

$$\frac{n_A-10000\times 0.5}{\sqrt{10000\times 0.5\times 0.5}}=\frac{n_A-5000}{50} \overset{\text{近似地}}{\sim} N(0,1).$$

从而

$$P\{n_A\geqslant 5325\}=P\left\{\frac{n_A-5000}{50}\geqslant\frac{5325-5000}{50}\right\}=1-P\left\{\frac{n_A-5000}{50}\leqslant 6.5\right\}\approx 1-\Phi(6.5)\approx 0.$$

即在硬币均匀的情况下, 出现“正面”5325 次几乎是不可能的.

故可断言硬币是不均匀的.

B9. 设随机变量 X 服从辛普森 (Simpson) 分布 (亦称三角分布), 密度函数为

$$f(x)=\begin{cases} x, & 0\leqslant x<1, \\ 2-x, & 1\leqslant x<2, \\ 0, & \text{其他}. \end{cases}$$

(1) 对 X 进行 100 次独立观察, 事件 $\{0.95<X<1.05\}$ 出现的次数记为 Y, 试用三种方法 (Y 的精确分布、用泊松分布来作为 Y 的近似分布、中心极限定理) 分别求 $P\{Y>2\}$;

(2) 要保证至少有 95% 的把握使事件 $\left\{\frac{1}{2}<X<\frac{3}{2}\right\}$ 出现的次数不少于 80, 问至少需要进行多少次观察?

解: (1) 由题可得

$$\begin{aligned} P\{0.95<X<1.05\} &= \int_{0.95}^{1} x\mathrm{d}x+\int_{1}^{1.05}(2-x)\mathrm{d}x=\frac{1}{2}x^2\Big|_{0.95}^{1}+\left(2x-\frac{1}{2}x^2\right)\Big|_{1}^{1.05} \\ &= \frac{1}{2}(1-0.95^2)+[(2\times 1.05-0.5\times 1.05^2)-(2-0.5)] \\ &= 0.04875+2.1-0.55125-1.5=0.0975. \end{aligned}$$

因此 $Y\sim B(100,0.0975)$.

(i) 利用 Y 的精确分布求解.

此时 $P\{Y>2\}=1-P\{Y\leqslant 2\}=1-(P\{Y=0\}+P\{Y=1\}+P\{Y=2\})$.

因为 $P\{Y=0\}=(1-0.0975)^{100}$, $P\{Y=1\}=\mathrm{C}_{100}^{1}(1-0.0975)^{99}\times 0.0975=9.75\times(0.9025)^{99}$;

$P\{Y=2\}=\mathrm{C}_{100}^{2}(1-0.0975)^{98}\times(0.0975)^2$.

从而 $P\{Y>2\}=1-(0.9025)^{98}[(0.9025)^2+9.75\times 0.9025+(0.0975)^2\times 4950]\approx 0.9976$.

(ii) 用泊松分布作 Y 近似分布求解.

此时取 $\lambda=100\times 0.0975=9.75$, 则此时 $P\{Y=k\}\approx\dfrac{\mathrm{e}^{-\lambda}\lambda^k}{k!}$.

从而

$$\begin{aligned} P\{Y>2\} &= 1-P\{Y\leqslant 2\}=1-(P\{Y=0\}+P\{Y=1\}+P\{Y=2\}) \\ &\approx 1-\left(\mathrm{e}^{-9.75}+9.75\mathrm{e}^{-9.75}+\frac{\mathrm{e}^{-9.75}9.75^2}{2!}\right) \\ &= 1-\mathrm{e}^{-9.75}(1+9.75+47.53125)\approx 0.9966. \end{aligned}$$

(iii) 用中心极限定理求解.

由中心极限定理

$$\frac{Y-100\times 0.0975}{\sqrt{100\times 0.0975\times(1-0.0975)}}\approx\frac{Y-9.75}{2.966}\overset{\text{近似地}}{\sim}N(0,1).$$

因此

$$P\{Y>2\}=1-P\{Y\leqslant 2\}=1-P\left\{\frac{Y-9.75}{2.966}\leqslant\frac{2-9.75}{2.966}\right\}\approx 1-\Phi(-2.61)=\Phi(2.61)\approx 0.9955.$$

(2) 设观察 n 次时事件 $\left\{\frac{1}{2}<X<\frac{3}{2}\right\}$ 发生的次数为 Z.

由于

$$\begin{aligned}P\left\{\frac{1}{2}<X<\frac{3}{2}\right\}&=\int_{\frac{1}{2}}^{1}x\mathrm{d}x+\int_{1}^{\frac{3}{2}}(2-x)\mathrm{d}x=\frac{1}{2}x^2\bigg|_{0.5}^{1}+\left(2x-\frac{1}{2}x^2\right)\bigg|_{1}^{1.5}\\&=\frac{1}{2}(1-0.5^2)+[(2\times 1.5-0.5\times 1.5^2)-(2-0.5)]\\&=0.375+3-1.125-1.5=0.75=\frac{3}{4}.\end{aligned}$$

从而此时有 $Z\sim B(n,0.75)$.

由中心极限定理

$$\frac{Z-0.75n}{\sqrt{0.75(1-0.75)n}}=\frac{Z-0.75n}{\sqrt{0.1875n}}\overset{\text{近似地}}{\sim}N(0,1).$$

由题 $P\{Z\geqslant 80\}=1-P\{Z<80\}=1-P\left\{\frac{Z-0.75n}{\sqrt{0.1875n}}<\frac{80-0.75n}{\sqrt{0.1875n}}\right\}\approx 1-\Phi\left(\frac{80-0.75n}{\sqrt{0.1875n}}\right)$.

至少 95% 的把握则有 $P\{Z\geqslant 80\}\geqslant 95\%$, 此时

$$\Phi\left(\frac{80-0.75n}{\sqrt{0.1875n}}\right)\leqslant 0.05\leqslant\Phi(-1.64).$$

此时有 $\frac{80-0.75n}{\sqrt{0.1875n}}\leqslant -1.64$, $\frac{3}{4}n-80\geqslant 1.64\sqrt{\frac{3}{16}n}$, 取 $x=\sqrt{n}$, 则有 $\frac{3}{4}x^2-\frac{1.64\sqrt{3}}{4}x-80\geqslant 0$.

判别式为 $\frac{1.64^2\times 3}{16}+3\times 80\approx 240.5$, 解得 $x\geqslant\frac{0.41\sqrt{3}+\sqrt{240.5}}{1.5}$ 或 $x\leqslant\frac{0.41\sqrt{3}-\sqrt{240.5}}{1.5}$.

由于 $x>0$, 则 $n\geqslant\left(\frac{\sqrt{240.5}+0.41\sqrt{3}}{1.5}\right)^2\approx 116.9$.

从而整数 n 至少取 117, 故至少需进行 117 次观察.

B10. 某企业庆祝百年华诞, 邀请了一些社会名流及企业的相关人士来参加庆典, 被邀请人独自一人或携伴 (一位同伴) 出席, 也有可能因故缺席, 这三种情况出现的概率分别为 0.3, 0.5, 0.2. 若此次庆典事先发出了 800 份邀请函, 若每位被邀请人参加庆典的行为相互独立, 问有超过千人出席该庆典的可能性有多大?

解: 由题, 设每份邀请函对应的出席人数为 $X_i, 1\leq i\leq 800$. 且 X_i 独立同分布.

考虑 X_1, 其分布律为 $P\{X_1=0\}=0.2$, $P\{X_1=1\}=0.3$, $P\{X_1=2\}=0.5$.

故 $E(X_1)=0\times 0.2+1\times 0.3+2\times 0.5=1.3$.

$\mathrm{Var}(X_1) = (0-1.3)^2 \times 0.2 + (1-1.3)^2 \times 0.3 + (2-1.3)^2 \times 0.5 = 0.338 + 0.027 + 0.245 = 0.61.$

由中心极限定理,

$$\frac{\frac{1}{800}\sum_{i=1}^{800} X_i - 1.3}{\sqrt{\frac{0.61}{800}}} = \frac{\sum_{i=1}^{800} X_i - 1040}{\sqrt{488}} \overset{\text{近似地}}{\sim} N(0,1).$$

超过千人出席即事件 $\left\{\sum_{i=1}^{800} X_i > 1000\right\}$. 因此

$$P\left\{\sum_{i=1}^{800} X_i > 1000\right\} = P\left\{\frac{\sum_{i=1}^{800} X_i - 1040}{\sqrt{488}} > \frac{1000-1040}{\sqrt{488}}\right\} = P\left\{\frac{\sum_{i=1}^{800} X_i - 1040}{\sqrt{488}} < -\frac{1000-1040}{\sqrt{488}}\right\}$$

$$\approx \Phi\left(\frac{20}{\sqrt{122}}\right) \approx \Phi(1.81) = 0.9649.$$

B11. 某次 "知识竞赛" 规则如下: 参赛选手最多可抽取 3 个相互独立的问题一一回答: 如果答错就被淘汰, 进而失去回答下一题的资格; 每答对一题得 1 分, 若 3 题都对则再加 1 分 (即共得 4 分). 现有 100 名参赛选手, 每人独立答题.

(1) 若每人至少答对一题的概率为 0.7, 用中心极限定理计算 "最多有 35 人得 0 分" 的概率;

(2) 若题目的难易程度类似, 每人答对每题的概率均为 0.8, 求这 100 名参赛选手的总分超过 220 分的概率.

解: (1) 由于每人至少答对一题概率为 0.7, 则一道题没答对, 即得 0 分的概率为 0.3.

设得 0 分的人数为 n_A, 则 $n_A \sim B(100, 0.3)$. 由中心极限定理,

$$\frac{n_A - 100 \times 0.3}{\sqrt{100 \times 0.3 \times 0.7}} = \frac{n_A - 30}{\sqrt{21}} \overset{\text{近似地}}{\sim} N(0,1).$$

从而 $P\{n_A \leqslant 35\} = P\left\{\frac{n_A - 30}{\sqrt{21}} \leqslant \frac{35-30}{\sqrt{21}}\right\} \approx \Phi\left(\frac{5}{\sqrt{21}}\right) \approx \Phi(1.09) = 0.8621.$

(2) 设第 i 人的得分为 $X_i, i = 1, 2, \cdots, 100$. 从而 X_i 独立同分布.

考虑 X_1, 当 $X_1 = 0$ 时即第一题就错了, $P\{X_1 = 0\} = 1 - 0.8 = 0.2$;

$X_1 = 1$ 时, 第一题对第二题错, $P\{X_1 = 1\} = 0.8 \times 0.2 = 0.16$;

$X_1 = 2$ 时, 前两题对第三题错, $P\{X_1 = 2\} = 0.8 \times 0.8 \times 0.2 = 0.128$;

$X_1 = 4$ 时, 三题都对, $P\{X_1 = 4\} = 0.8^3 = 0.512$.

从而 $E(X_1) = 0 \times 0.2 + 1 \times 0.16 + 2 \times 0.128 + 4 \times 0.512 = 2.464.$

$E(X_1^2) = 0^2 \times 0.2 + 1^2 \times 0.16 + 2^2 \times 0.128 + 4^2 \times 0.512 = 8.864;$

则 $\mathrm{Var}(X_1) = E(X_1^2) - (E(X_1))^2 = 2.792704 \approx 2.793.$

由中心极限定理

$$\frac{\sum_{i=1}^{100} X_i - 2.464 \times 100}{\sqrt{2.793 \times 100}} = \frac{\sum_{i=1}^{100} X_i - 246.4}{\sqrt{279.3}} \overset{\text{近似地}}{\sim} N(0,1).$$

从而

$$P\left\{\sum_{i=1}^{100} X_i > 220\right\} = P\left\{\frac{\sum_{i=1}^{100} X_i - 246.4}{\sqrt{279.3}} > \frac{220 - 246.4}{\sqrt{279.3}}\right\} = P\left\{\frac{\sum_{i=1}^{100} X_i - 246.4}{\sqrt{279.3}} < -\frac{220 - 246.4}{\sqrt{279.3}}\right\}$$
$$\approx \Phi\left(\frac{26.4}{\sqrt{279.3}}\right) \approx \Phi(1.58) = 0.9429.$$

第 6 章　统计量与抽样分布

A1. 假设 $X_1, X_2, \cdots, X_n$ 是从总体 X 中抽取的样本 $(n \geqslant 1)$. 当总体 X 服从如下分布时, 写出样本的联合分布律或联合密度函数:

(1) 总体服从二项分布 $B(10, 0.2)$;

(2) 总体服从泊松分布 $P(1)$;

(3) 总体服从标准正态分布 $N(0, 1)$;

(4) 总体服从指数分布 $E(1)$.

解: (1) 此时 $P\{X_i = x\} = \mathrm{C}_{10}^{x} 0.2^x (1-0.2)^{10-x}$, $x = 0, 1, \cdots, 10$.

从而联合分布律为

$$p(x_1, x_2, \cdots, x_n) = \prod_{i=1}^{n} P\{X_i = x_i\} = \prod_{i=1}^{n} \mathrm{C}_{10}^{x_i} 0.2^{x_i} 0.8^{10-x_i}, x_i = 0, 1, \cdots, 10, i = 1, 2, \cdots, n.$$

(2) 此时 $P\{X_i = k\} = \dfrac{e^{-1} \cdot 1^k}{k!} = \dfrac{e^{-1}}{k!}$, $k = 0, 1, 2, \cdots$.

从而联合分布律为

$$p(x_1, x_2, \cdots, x_n) = \prod_{i=1}^{n} P\{X_i = x_i\} = \prod_{i=1}^{n} \frac{e^{-1}}{x_i!}, x_i = 0, 1, \cdots, i = 1, 2, \cdots, n.$$

(3) 此时 $f_{X_i}(x) = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$, $x \in \mathbb{R}$.

$$f(x_1, x_2, \cdots, x_n) = \prod_{i=1}^{n} f_{X_i}(x_i) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} e^{-\frac{x_i^2}{2}}, x_i \in \mathbb{R}, i = 1, 2, \cdots, n.$$

(4) 此时 $f_{X_i}(x) = e^{-x}$, $x > 0$.

$$f(x_1, x_2, \cdots, x_n) = \prod_{i=1}^{n} P\{X_i = x_i\} = \prod_{i=1}^{n} e^{-x_i}, x_i > 0, i = 1, 2, \cdots, n.$$

A2. 设总体 $X \sim N(\mu, 1)$, μ 未知, $X_1, X_2, \ldots, X_5$ 是来自总体 X 的简单随机样本, 判断下列哪些是统计量, 哪些不是统计量:

(1) $\sum\limits_{i=1}^{5} X_i$;　　(2) $\sum\limits_{i=1}^{5} X_i^2 - 5\mu^2$;

(3) $\sum\limits_{i=1}^{5} (X_i - \mu)$;　　(4) $X_1 - X_2$.

解: 由于(1)(4)中不包含未知量, 则(1)(4)为统计量.

因为(2)(3)中包含未知量 μ, 因此(2)(3)不是统计量.

A3. 从总体 X 中抽取样本容量为 5 的样本, 其观测值为 2.6, 4.1, 3.2, 3.6,2.9, 计算样本均值、样本方差和样本二阶中心矩.

解: 样本均值

$$\overline{X}=\frac{1}{5}(2.6+4.1+3.2+3.6+2.9)=3.28.$$

样本方差

$$S^2=\frac{1}{4}[(2.6-3.28)^2+(4.1-3.28)^2+(3.2-3.28)^2+(3.6-3.28)^2+(2.9-3.28)^2]=0.347.$$

样本二阶中心矩

$$B_2=\frac{1}{5}[(2.6-3.28)^2+(4.1-3.28)^2+(3.2-3.28)^2+(3.6-3.28)^2+(2.9-3.28)^2]=0.2776.$$

A4. 从总体 $X\sim N(0,1)$ 中抽取样本容量为 10 的样本, 其观测值为 2.50, 0.49, 0.53, -0.37, 0.61, -0.63, 0.01, 0.81, 0.78, 0.27. 计算样本均值和样本方差.

解: 样本均值

$$\overline{X}=\frac{1}{10}(2.50+0.49+0.53-0.37+0.61-0.63+0.01+0.81+0.78+0.27)=0.5.$$

样本方差

$$\begin{aligned}S^2=\frac{1}{9}[&(2.50-0.5)^2+(0.49-0.5)^2+(0.53-0.5)^2+(-0.37-0.5)^2+(0.61-0.5)^2+(-0.63-0.5)^2\\&+(0.01-0.5)^2+(0.81-0.5)^2+(0.78-0.5)^2+(0.27-0.5)^2]=0.7238.\end{aligned}$$

A5. 假设 $X_1,X_2,\cdots,X_7$ 是从总体 $X\sim B(1,0.3)$ 中抽取的简单随机样本.

(1) 求样本均值 $\overline{X}$ 的数学期望和方差;

(2) 求样本方差 S^2 的数学期望;

(3) 求 $P\{\max\{X_1,X_2,\cdots,X_7\}<1\}$.

解: (1) 由题, $\sum\limits_{i=1}^{7}X_i\sim B(7,0.3)$.

则 $E\left(\sum\limits_{i=1}^{7}X_i\right)=7\times0.3=2.1$, $\mathrm{Var}\left(\sum\limits_{i=1}^{7}X_i\right)=7\times0.3\times(1-0.3)=1.47$.

又因为 $\overline{X}=\frac{1}{7}\sum\limits_{i=1}^{7}X_i$, 则 $E(\overline{X})=\frac{1}{7}E\left(\sum\limits_{i=1}^{7}X_i\right)=0.3$; $\mathrm{Var}\ (\overline{X})=\frac{1}{7^2}\mathrm{Var}\left(\sum\limits_{i=1}^{7}X_i\right)=0.03$.

(2) 由于 $S^2 = \frac{1}{6}\sum_{i=1}^{7}(X_i - \overline{X})^2 = \frac{1}{6}\left(\sum_{i=1}^{7} X_i^2 - 7\overline{X}^2\right)$.

因为 $E(X_1^2) = \mathrm{Var}(X_1) + (E(X_1))^2 = 1 \times 0.3 \times 0.7 + (1 \times 0.3)^2 = 0.3$.

$E(\overline{X}^2) = \mathrm{Var}(\overline{X}) + (E(\overline{X}))^2 = 0.03 + (0.3)^2 = 0.12$.

因此

$$E(S^2) = \frac{1}{6}E\left(\sum_{i=1}^{7} X_i^2 - 7\overline{X}^2\right) = \frac{7}{6}(E(X_1^2) - E(\overline{X}^2)) = \frac{7}{6}(0.3 - 0.12) = 0.21.$$

(3)

$$\begin{aligned} P\{\max\{X_1, X_2, \cdots, X_7\} < 1\} &= P\{X_1 < 1, X_2 < 1, \cdots, X_7 < 1\} \\ &= \prod_{i=1}^{7} P\{X_i < 1\} = \prod_{i=1}^{7} P\{X_i = 0\} = (1 - 0.3)^7 = 0.7^7. \end{aligned}$$

A6. 给出下列上分位数的值:

(1) $\chi^2_{0.05}(5), \chi^2_{0.06}(5), \chi^2_{0.95}(5), \chi^2_{0.94}(5)$;

(2) $t_{0.05}(8), t_{0.06}(8), t_{0.95}(8), t_{0.94}(8)$;

(3) $F_{0.05}(3,5), F_{0.05}(5,3), F_{0.04}(3,5), F_{0.04}(5,3)$.

解: 查书后表, 以及利用 Excel 计算可得.

(1) $\chi^2_{0.05}(5) = 11.070$, $\chi^2_{0.06}(5) = 10.5962$, $\chi^2_{0.95}(5) = 1.145$, $\chi^2_{0.94}(5) = 1.2499$.

(2) $t_{0.05}(8) = 1.8595$, $t_{0.06}(8) = 1.7402$, $t_{0.95}(8) = -1.8595$, $t_{0.94}(8) = -1.7402$.

(3) $F_{0.05}(3,5) = 5.41$, $F_{0.05}(5,3) = 9.01$, $F_{0.04}(3,5) = 6.098$, $F_{0.04}(5,3) = 10.617$.

A7. 假设 $X_1, X_2, \cdots, X_5$ 是从总体 $X \sim \chi^2(2)$ 中抽取的简单随机样本.

(1) 求 $P\{X_1 + X_2 + \cdots + X_5 > 18.307\}$;

(2) 求 $X_1 + X_2 + \cdots + X_5$ 的分布, 并由此给出它的上 0.1 分位数.

解: (1) 由简单随机样本可得, $X_1 + X_2 + X_3 + X_4 + X_5 \sim \chi^2(10)$.

又查表知 $\chi^2_{0.05}(10) = 18.307$, 则 $P\{X_1 + X_2 + X_3 + X_4 + X_5 > 18.307\} = 0.05$.

(2) 由(1), $X_1 + X_2 + X_3 + X_4 + X_5 \sim \chi^2(10)$.

查表知它的上 0.1 分位数为 $\chi^2_{0.1}(10) = 15.987$.

A8. 假设 $X \sim N(0,1)$, 利用 χ^2 分布的性质, 求 X 的四阶原点矩 A_4.

解: 由题, $Y = X^2 \sim \chi^2(1)$.

由 χ^2 分布性质, $E(Y) = 1$, $\mathrm{Var}(Y) = 2$. 从而 $E(Y^2) = \mathrm{Var}(Y) + (E(Y))^2 = 2 + 1 = 3$.

因此 $A_4 = E(X^4) = E(Y^2) = 3$.

A9. 假设 $X_1, X_2, \cdots, X_{10}$ 为来自总体 $X \sim N(0,1)$ 的一个简单随机样本, 记 $Y^2 = X_1^2 + X_2^2 + \cdots + X_9^2$, $T = \dfrac{3X_{10}}{Y}$.

(1) 求 $P\{|T| > 1.8331\}$;

(2) 求 T 的上 0.10 分位数.

解: 由题, $X_i^2 \sim \chi^2(1)$, $i = 1, 2, \cdots, 9$. 则由独立可知 $Y^2 \sim \chi^2(9)$.

又因为 $X_{10} \sim N(0,1)$, 则 $T = \dfrac{3X_{10}}{Y} = \dfrac{X_{10}}{\sqrt{\dfrac{Y^2}{9}}} \sim t(9)$.

因此 $P\{|T| > 1.8331\} = 2P\{T > 1.8331\} = 2P\{T > t_{0.05}(9)\} = 2 \times 0.05 = 0.1$.

(2) 查表知 T 的上 0.10 分位数即为 $t_{0.1}(9) = 1.3830$.

A10. 假设 $X \sim t(5)$, 求 $Y = \dfrac{1}{X^2}$ 的上 0.05 分位数和上 0.1 分位数.

解: 由 t 分布定义, 将 $X \sim t(5)$ 表示为 $X = \dfrac{U}{\sqrt{\dfrac{V}{5}}}$, 其中 $U \sim N(0,1)$, $V \sim \chi^2(5)$.

又因为 $U^2 \sim \chi^2(1)$, 从而 $Y = \dfrac{1}{X^2} = \dfrac{\frac{1}{5}V}{U^2} \sim F(5,1)$.

从而查表得 Y 的上 0.05 分位数为 $F_{0.05}(5,1) = 230$, 上 0.1 分位数 $F_{0.1}(5,1) = 57.24$.

A11. 设总体 X 服从标准正态分布, $X_1, X_2, \cdots, X_{16}$ 是来自总体 X 的简单随机样本, 写出下列统计量的分布:

(1) 样本均值 $\overline{X}$; (2) $\sum\limits_{i=1}^{16} X_i^2$; (3) $\dfrac{3X_1}{\sqrt{\sum\limits_{i=2}^{10} X_i^2}}$; (4) $\dfrac{X_1 + X_2}{\sqrt{X_3^2 + X_4^2}}$; (5) $\overline{X} - X_1$.

解: (1) 因为 $X_i \sim N(0,1)$, $i = 1, 2, \cdots, 16$, 且相互独立, 则 $\sum\limits_{i=1}^{16} X_i \sim N(0, 16)$.

因此 $\overline{X} = \dfrac{1}{16}\sum\limits_{i=1}^{16} X_i \sim N\left(0, \dfrac{1}{16}\right)$.

(2) 因为 $X_i^2 \sim \chi^2(1)$, $i = 1, 2, \cdots, 16$, 且相互独立, 则 $\sum\limits_{i=1}^{16} X_i^2 \sim \chi^2(16)$.

(3) 因为 $\sum\limits_{i=2}^{10} X_i^2 \sim \chi^2(9)$. 则 $\dfrac{3X_1}{\sqrt{\sum\limits_{i=2}^{10} X_i^2}} = \dfrac{X_1}{\sqrt{\dfrac{1}{9}\sum\limits_{i=2}^{10} X_i^2}} \sim t(9)$.

(4) 因为 $X_3^2 + X_4^2 \sim \chi^2(2)$, 且 $X_1 + X_2 \sim N(0, 2)$, 从而

$$\frac{X_1 + X_2}{\sqrt{X_3^2 + X_4^2}} = \frac{\dfrac{(X_1 + X_2)}{\sqrt{2}}}{\sqrt{\dfrac{X_3^2 + X_4^2}{2}}} \sim t(2).$$

(5) $\overline{X}-X_1=\frac{1}{16}\sum_{i=1}^{16}X_i-X_1=-\frac{15}{16}X_1+\frac{1}{16}\sum_{i=2}^{16}X_i$.

此时 $\sum_{i=2}^{16}X_i\sim N(0,15)$, 则 $\frac{1}{16}\sum_{i=2}^{16}X_i\sim N\left(0,\frac{15}{16^2}\right)$, 又因为 $-\frac{15}{16}X_1\sim N\left(0,\frac{15^2}{16^2}\right)$.

因此

$$\overline{X}-X_1\sim N\left(0,\frac{15^2}{16^2}+\frac{15}{16^2}\right)=N\left(0,\frac{15}{16}\right).$$

A12. 假设总体 $X\sim N(0,4),X_1,X_2,\ldots,X_{20}$ 是来自总体 X 的简单随机样本, 求 $P\left\{33.04\leqslant\sum_{i=1}^{20}X_i^2\leqslant 125.64\right\}$.

解: 由题, 因为 $\frac{X_i}{2}\sim N(0,1)$, 从而 $\frac{X_i^2}{4}\sim\chi^2(1)$, $i=1,2,\cdots,20$.

因此 $\frac{1}{4}\sum_{i=1}^{20}X_i^2\sim\chi^2(20)$. 从而

$$P\left\{33.04\leqslant\sum_{i=1}^{20}X_i^2\leqslant 125.64\right\}=P\left\{8.26\leqslant\frac{1}{4}\sum_{i=1}^{20}X_i^2\leqslant 31.41\right\}=P\left\{\chi^2_{0.99}(20)\leqslant\frac{1}{4}\sum_{i=1}^{20}X_i^2\leqslant\chi^2_{0.05}(20)\right\}$$

$$=0.99-0.05=0.94.$$

A13. 设总体 $X\sim N(\mu,\sigma^2)$, $X_1,X_2,\cdots,X_{25}$ 是来自总体 X 的简单随机样本, μ 和 σ^2 均未知, S^2 为样本方差, 求 $P\left\{\frac{S^2}{\sigma^2}\leqslant 0.577\right\}$.

解: 由题, $\frac{24S^2}{\sigma^2}\sim\chi^2(24)$. 从而有

$$P\left\{\frac{S^2}{\sigma^2}\leqslant 0.577\right\}=P\left\{\frac{24S^2}{\sigma^2}\leqslant 13.848\right\}=1-P\left\{\frac{24S^2}{\sigma^2}>\chi^2_{0.95}(24)\right\}=1-0.95=0.05.$$

B1. 设总体 $X\sim N(\mu,\sigma^2)$, $X_1,X_2,\ldots,X_{25}$ 是来自总体 X 的简单随机样本, $\overline{X}$ 是样本均值.

(1)求 $P\{|\overline{X}-\mu|<0.2\sigma\}$;

(2)若 $P\{\overline{X}>\mu-c\sigma\}=0.95$, 求 c.

解: (1) 由题 $\overline{X}\sim N\left(\mu,\dfrac{\sigma^2}{25}\right)$, 则此时 $\dfrac{\overline{X}-\mu}{\frac{\sigma}{5}}\sim N(0,1)$.

因此

$$P\{|\overline{X}-\mu|<0.2\sigma\}=P\left\{\left|\frac{\overline{X}-\mu}{\frac{\sigma}{5}}\right|<1\right\}=\Phi(1)-\Phi(-1)=2\Phi(1)-1=0.6826.$$

(2) 因为

$$P\{\overline{X}>\mu-c\sigma\}=1-P\left\{\frac{\overline{X}-\mu}{\frac{\sigma}{5}}\leqslant -5c\right\}=1-\Phi(-5c)=\Phi(5c)=0.95.$$

从而查表得 $5c=1.65$, 则 $c=0.33$.

B2. 设总体 $X\sim N(\mu,\sigma^2), X_1,X_2,\cdots,X_9$ 是来自总体 X 的简单随机样本, $\overline{X}$ 是样本均值, S^2 是样本方差, 写出下列抽样分布:

(1) $\dfrac{3(\overline{X}-\mu)}{\sigma}$; (2) $\dfrac{3(\overline{X}-\mu)}{S}$;

(3) $\dfrac{\sum\limits_{i=1}^{9}(X_i-\overline{X})^2}{\sigma^2}$; (4) $\dfrac{\sum\limits_{i=1}^{9}(X_i-\mu)^2}{\sigma^2}$;

(5) $\dfrac{9(\overline{X}-\mu)^2}{\sigma^2}$; (6) $\dfrac{9(\overline{X}-\mu)^2}{S^2}$;

(7) $\dfrac{2(X_1-X_2)^2}{(X_3-X_4)^2+(X_5-X_6)^2}$;

(8) $\dfrac{(X_1-Y_1)^2+(X_2-Y_1)^2+(X_3-Y_1)^2}{(X_4-Y_2)^2+(X_5-Y_2)^2+(X_6-Y_2)^2}$, 其中 $Y_1=\dfrac{X_1+X_2+X_3}{3}$, $Y_2=\dfrac{X_4+X_5+X_6}{3}$.

解: 由题, $\overline{X}\sim N\left(\mu,\dfrac{\sigma^2}{9}\right)$, $\dfrac{8S^2}{\sigma^2}\sim\chi^2(8)$.

(1) 因为 $\overline{X}-\mu\sim N\left(0,\dfrac{\sigma^2}{9}\right)$, 则 $\dfrac{3(\overline{X}-\mu)}{\sigma}\sim N(0,1)$.

(2) 由(1)中结果有

$$\frac{3(\overline{X}-\mu)}{S}=\frac{\frac{3(\overline{X}-\mu)}{\sigma}}{\sqrt{\frac{1}{8}\cdot\frac{8S^2}{\sigma^2}}}\sim t(8).$$

(3) 由题 $\dfrac{\sum\limits_{i=1}^{9}(X_i-\overline{X})^2}{\sigma^2}=\dfrac{8S^2}{\sigma^2}\sim\chi^2(8)$.

(4) 由于 $\dfrac{X_i-\mu}{\sigma}\sim N(0,1)$, $i=1,2,\cdots,9$. 因此

$$\frac{\sum\limits_{i=1}^{9}(X_i-\mu)^2}{\sigma^2}=\sum_{i=1}^{9}\left(\frac{X_i-\mu}{\sigma}\right)^2\sim\chi^2(9).$$

(5) 由(1)可知 $\dfrac{9(\overline{X}-\mu)^2}{\sigma^2}=\left[\dfrac{3(\overline{X}-\mu)}{\sigma}\right]^2\sim\chi^2(1)$.

(6) 由(2)并结合 F 分布的性质, $\dfrac{9(\overline{X}-\mu)^2}{S^2}=\left[\dfrac{3(\overline{X}-\mu)}{S}\right]^2\sim F(1,8)$.

(7) 因为 X_1-X_2, X_3-X_4, X_5-X_6 均服从 $N(0,2\sigma^2)$, 且相互独立.

从而 $\dfrac{(X_1-X_2)^2}{2\sigma^2}$, $\dfrac{(X_3-X_4)^2}{2\sigma^2}$, $\dfrac{(X_5-X_6)^2}{2\sigma^2}$ 均服从 $\chi^2(1)$.

此时 $\dfrac{(X_3-X_4)^2+(X_5-X_6)^2}{2\sigma^2}\sim\chi^2(2)$, 则有

$$\frac{2(X_1-X_2)^2}{(X_3-X_4)^2+(X_5-X_6)^2}=\frac{\dfrac{2(X_1-X_2)^2}{2\sigma^2}}{\dfrac{(X_3-X_4)^2+(X_5-X_6)^2}{2\sigma^2}}=\frac{\dfrac{(X_1-X_2)^2}{2\sigma^2}}{\dfrac{1}{2}\cdot\dfrac{(X_3-X_4)^2+(X_5-X_6)^2}{2\sigma^2}}\sim F(1,2).$$

(8) 由于 Y_1 为 X_1,X_2,X_3 的样本均值, 故样本方差 $S_1^2=\dfrac{1}{2}[(X_1-Y_1)^2+(X_2-Y_1)^2+(X_3-Y_1)^2]$.

且此时 $\dfrac{2S_1^2}{\sigma^2}\sim\chi^2(2)$.

类似地, 对 $S_2^2=\dfrac{1}{2}[(X_4-Y_2)^2+(X_5-Y_2)^2+(X_6-Y_2)^2]$, $\dfrac{2S_2^2}{\sigma^2}\sim\chi^2(2)$.

从而

$$\frac{(X_1-Y_1)^2+(X_2-Y_1)^2+(X_3-Y_1)^2}{(X_4-Y_2)^2+(X_5-Y_2)^2+(X_6-Y_2)^2}=\frac{S_1^2}{S_2^2}=\frac{\dfrac{2S_1^2}{\sigma^2}\cdot\dfrac{1}{2}}{\dfrac{2S_2^2}{\sigma^2}\cdot\dfrac{1}{2}}\sim F(2,2).$$

B3. 假设二维总体 $(X,Y)\sim N(0,0;1,1;\rho)$, (X_i,Y_i), $i=1,2,\cdots,10$ 为从该总体中抽取的简单随机样本, 记统计量 $Z=a\sum\limits_{i=1}^{10}(X_i+Y_i)^2$, 若 $Z\sim\chi^2(n)$, 求 a 和 n 的值.

解: 由题, $X\sim N(0,1)$, $Y\sim N(0,1)$, 且 $\rho_{XY}=\dfrac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}=\rho$, 从而 $\mathrm{Cov}(X,Y)=\rho$.

从而 $E(X_i+Y_i)=E(X_i)+E(Y_i)=0$,

$\mathrm{Var}(X_i+Y_i)=\mathrm{Var}(X_i)+\mathrm{Var}(Y_i)+2\mathrm{Cov}(X_i,Y_i)=1+1+2\rho=2(1+\rho)$.

由线性组合则 X_i+Y_i 服从正态分布, 此时 $X_i+Y_i\sim N(0,2(1+\rho))$, $i=1,2,\cdots,10$.

从而 $\dfrac{X_i+Y_i}{\sqrt{2(1+\rho)}}\sim N(0,1)$, 因此 $\dfrac{(X_i+Y_i)^2}{2(1+\rho)}\sim\chi^2(1)$, 则 $\sum\limits_{i=1}^{10}\dfrac{(X_i+Y_i)^2}{2(1+\rho)}\sim\chi^2(10)$.

因为 $Z=a\sum\limits_{i=1}^{10}(X_i+Y_i)^2\sim\chi^2(10)$, 对比系数有 $a=\dfrac{1}{2(1+\rho)}$, $n=10$.

B4. 对一重量为 a 的物体独立重复称 n 次, 现准备用这 n 次读数的平均值去估计 a. 假设这批读数来自均值为 a, 标准差为 2.5 的正态总体, 至少要称多少次才能使估计值与 a 之差的绝对值不大于 0.5 的概率

(1) 超过 90%; (2) 超过 95% .

解: 设第 i 次称的读数为 X_i, 则 $X_i \sim N(a, 2.5^2)$, $i = 1, 2, \cdots, n$. 且 X_i 间相互独立.

此时 $\overline{X} = \dfrac{1}{n}\sum\limits_{i=1}^{n} X_i \sim N\left(a, \dfrac{2.5^2}{n}\right)$. 且所求概率为 $P\{|\overline{X} - a| \leqslant 0.5\}$.

且由于 $\dfrac{\overline{X} - a}{\dfrac{2.5}{\sqrt{n}}} \sim N(0, 1)$, 则

$$P\{|\overline{X} - a| \leqslant 0.5\} = P\left\{\left|\frac{\overline{X} - a}{\dfrac{2.5}{\sqrt{n}}}\right| \leqslant \frac{0.5\sqrt{n}}{2.5}\right\} = \Phi(0.2\sqrt{n}) - \Phi(-0.2\sqrt{n}) = 2\Phi(0.2\sqrt{n}) - 1.$$

(1) 当 $2\Phi(0.2\sqrt{n}) - 1 > 0.9$ 时, $\Phi(0.2\sqrt{n}) > 0.95$, 查表得 $0.2\sqrt{n} > 1.645$,

有 $n > 67.65$, 从而至少称 68 次.

(2) 当 $2\Phi(0.2\sqrt{n}) - 1 > 0.95$ 时, $\Phi(0.2\sqrt{n}) > 0.975$, 查表得 $0.2\sqrt{n} > 1.96$,

有 $n > 96.04$, 从而至少称 97 次.

B5. 设总体 X 的密度函数

$$f(x) = \frac{1}{2}\mathrm{e}^{-|x|}, \quad -\infty < x < +\infty$$

从总体中抽取样本容量为 10 的样本, $\overline{X}$ 和 S^2 分别是样本均值和样本方差. 求:

(1) $\overline{X}$ 的数学期望和方差;

(2) S^2 的数学期望.

解: 由第四章 B20 题的结果, $E(X) = 0$, $E(X^2) = 2$, $\mathrm{Var}(X) = 2$.

(1) 因为 $E(X_i) = 0$, $\mathrm{Var}(X_i) = 2$, $E(X_i^2) = 2$, $i = 1, 2, \cdots, 10$, 从而

$$E(\overline{X}) = E\left(\frac{1}{10}\sum_{i=1}^{10} X_i\right) = \frac{1}{10}\sum_{i=1}^{10} E(X_i) = 0.$$

$$\mathrm{Var}(\overline{X}) = \mathrm{Var}\left(\frac{1}{10}\sum_{i=1}^{10} X_i\right) = \frac{1}{100}\sum_{i=1}^{10} \mathrm{Var}(X_i) = \frac{20}{100} = \frac{1}{5}.$$

(2)

$$E(S^2) = E\left(\frac{1}{9}\sum_{i=1}^{10}(X_i - \overline{X})^2\right) = \frac{1}{9}E\left(\sum_{i=1}^{10}(X_i^2 - 2X_i\overline{X} + \overline{X}^2)\right) = \frac{1}{9}E\left(\sum_{i=1}^{10} X_i^2 - 10\overline{X}^2\right)$$

$$= \frac{1}{9}\sum_{i=1}^{10} E(X_i^2) - \frac{10}{9}E(\overline{X}^2) = \frac{1}{9}\times 10\times 2 - \frac{10}{9}[\mathrm{Var}(\overline{X}) + (E(\overline{X}))^2]$$
$$= \frac{20}{9} - \frac{10}{9}\times\frac{1}{5} = 2.$$

【结合例题, 事实上有 $E(\overline{X}) = E(X)$, $\mathrm{Var}(\overline{X}) = \dfrac{\mathrm{Var}(X)}{n}$, $E(S^2) = \mathrm{Var}(X)$.】

【因此后续题目直接用上述结论来呈现】

B6. 设总体 $X \sim U(0,\theta)$, $X_1, X_2, \cdots, X_5$ 是来自总体 X 的简单随机样本, $\overline{X}$ 是样本均值, S^2 是样本方差, 求 $E(\overline{X}), E(\overline{X}^2)$ 和 $E(S^2)$.

解: 由均匀分布可知, $E(X) = \dfrac{\theta}{2}$, $\mathrm{Var}(X) = \dfrac{\theta^2}{12}$.

因此 $E(\overline{X}) = \dfrac{\theta}{2}$, $\mathrm{Var}(\overline{X}) = \dfrac{\theta^2}{12}\cdot\dfrac{1}{5} = \dfrac{\theta^2}{60}$.

从而 $E(\overline{X}^2) = \mathrm{Var}(\overline{X}) + (E(\overline{X}))^2 = \dfrac{\theta^2}{60} + \dfrac{\theta^2}{4} = \dfrac{16\theta^2}{60} = \dfrac{4\theta^2}{15}$.

且 $E(S^2) = \dfrac{\theta^2}{12}$.

B7. 设总体 X 的密度函数

$$f(x) = \begin{cases} \lambda \mathrm{e}^{-\lambda x}, & x > 0, \\ 0, & x \leqslant 0, \end{cases}$$

从总体中抽取样本容量为 10 的样本.

(1) 求样本均值的数学期望和方差;

(2) 记 $X_{(1)} = \min\{X_1, X_2, \cdots, X_{10}\}$, 求 $X_{(1)}$ 的数学期望和方差.

解: (1) 总体服从指数分布, 此时 $E(X) = \dfrac{1}{\lambda}$, $\mathrm{Var}(X) = \dfrac{1}{\lambda^2}$.

(1) 由题, $E(\overline{X}) = \dfrac{1}{\lambda}$, $\mathrm{Var}(\overline{X}) = \dfrac{1}{10\lambda^2}$.

(2) 设 $X_{(1)}$ 的分布函数为 $F_N(x)$, 从而 $F_N(x) = 1 - \prod\limits_{i=1}^{10}(1 - F_{X_i}(x))$,

又因为 $x > 0$ 时 $F_{X_i}(x) = 1 - e^{-\lambda x}$, $i = 1, 2, \cdots, 10$.

从而 $x > 0$ 时 $F_N(x) = 1 - \prod\limits_{i=1}^{10}[1 - (1 - e^{-\lambda x})] = 1 - e^{-10\lambda x}$.

又 $x \leqslant 0$ 时 $F_N(x) = 0$, 则此时密度函数 $f_N(x) = \begin{cases} 10e^{-10\lambda x}, & x > 0, \\ 0, & x \leqslant 0. \end{cases}$

因此 $X_{(1)} \sim E(10\lambda)$, 因此 $E(X_{(1)}) = \dfrac{1}{10\lambda}$, $\mathrm{Var}(X_{(1)}) = \dfrac{1}{100\lambda^2}$.

B8. 假设总体 X 服从指数分布, 密度函数为

$$f(x)=\begin{cases}\dfrac{1}{20}\mathrm{e}^{-\frac{1}{20}x}, & x>0,\\ 0, & x\leqslant 0,\end{cases}$$

从总体中抽取一个样本容量为 50 的简单随机样本 $X_1,X_2,\cdots,X_{50}$, 记 $Y_j=\min\limits_{1+10j\leqslant i\leqslant 10+10j}X_i$, $j=0,1,2,3,4$, 求 $Z=\sum\limits_{i=0}^{4}Y_j$ 的分布.

解: 类似 B7 的讨论, 因为 $X\sim E\left(\dfrac{1}{20}\right)$, 则 $Y_j\sim E\left(\dfrac{1}{2}\right)$, $j=0,1,2,3,4$.

【 Y_j 相当于每十个样本的最小值】

此时 Y_j 的密度函数为 $f_{Y_j}(x)=\begin{cases}\dfrac{1}{2}e^{-\frac{x}{2}}=\dfrac{1}{2^1\Gamma(1)}x^{1-1}e^{-\frac{x}{2}}, & x>0,\\ 0, & x\leqslant 0.\end{cases}$

从而 $Y_j\sim\chi^2(2)$, 因此 $Z=\sum\limits_{i=0}^{4}Y_j\sim\chi^2(10)$.

B9. 设 $X_1,X_2,\cdots,X_8$ 是来自标准正态总体的样本, $\overline{X}=\dfrac{1}{8}\sum\limits_{i=1}^{8}X_i$, $S^2=\dfrac{1}{7}\sum\limits_{i=1}^{8}(X_i-\overline{X})^2$, X_9 是新增的样本, 试确定 $Y=\dfrac{2\sqrt{2}}{3}\cdot\dfrac{X_9-\overline{X}}{S}$ 的分布.

解: 因为 $\overline{X}\sim N\left(0,\dfrac{1}{8}\right)$, $X_9\sim N(0,1)$, 则 $X_9-\overline{X}$ 服从正态分布.

且 $E(X_9-\overline{X})=E(X_9)-E(\overline{X})=0$.

$$\mathrm{Var}(X_9-\overline{X})=\mathrm{Var}\left(X_9-\frac{1}{8}\sum_{i=1}^{8}X_i\right)=\mathrm{Var}(X_9)+\frac{1}{64}\sum_{i=1}^{8}\mathrm{Var}(X_i)=1+\frac{1}{64}\times 8=\frac{9}{8}.$$

从而 $X_9-\overline{X}\sim N\left(0,\dfrac{9}{8}\right)$, 则 $\dfrac{X_9-\overline{X}}{\frac{3}{2\sqrt{2}}}=\dfrac{2\sqrt{2}}{3}(X_9-\overline{X})\sim N(0,1)$.

又因为 $\dfrac{7S^2}{1}=7S^2\sim\chi^2(7)$, 因此

$$Y=\frac{2\sqrt{2}}{3}\frac{X_9-\overline{X}}{S}=\frac{\frac{2\sqrt{2}}{3}(X_9-\overline{X})}{\sqrt{\frac{7S^2}{7}}}\sim t(7).$$

B10. 设总体 $X\sim N(\mu,\sigma^2)$, $X_1,X_2,\cdots,X_5$ 和 $Y_1,Y_2,\cdots,Y_9$ 是来自总体 X 的两个独立样本, $\overline{X}$ 和 $\overline{Y}$ 分别是两个样本的样本均值, S_1^2 和 S_2^2 分别是两个样本的样本方差.

(1) 若 $\dfrac{a(\overline{X}-\overline{Y})}{\sigma}\sim N(0,1)$, 求 a;

(2) 若 $\dfrac{b(\overline{X}-\overline{Y})}{\sqrt{S_1^2+2S_2^2}}\sim t(12)$, 求 b.

解: (1) 由题 $\overline{X}\sim N\left(\mu,\dfrac{\sigma^2}{5}\right)$, $\overline{Y}\sim N\left(\mu,\dfrac{\sigma^2}{9}\right)$. 且 $\overline{X}$ 与 $\overline{Y}$ 独立.

此时有 $\mathrm{Var}(\overline{X}-\overline{Y})=\mathrm{Var}(\overline{X})+\mathrm{Var}(\overline{Y})=\dfrac{14}{45}\sigma^2$. 从而 $\overline{X}-\overline{Y}\sim N\left(0,\dfrac{14}{45}\sigma^2\right)$.

因此 $\dfrac{a(\overline{X}-\overline{Y})}{\sigma}\sim N\left(0,\dfrac{14a^2}{45}\right)$, 由题则 $\dfrac{14a^2}{45}=1$, 解得 $a=\pm\sqrt{\dfrac{45}{14}}$.

(2) 因为 $\dfrac{4S_1^2}{\sigma^2}\sim\chi^2(4)$, $\dfrac{8S_2^2}{\sigma^2}\sim\chi^2(8)$. 因此有 $\dfrac{4S_1^2+8S_2^2}{\sigma^2}\sim\chi^2(12)$.

从而有

$$\frac{\dfrac{a(\overline{X}-\overline{Y})}{\sigma}}{\sqrt{\dfrac{4S_1^2+8S_2^2}{12\sigma^2}}}=\frac{\sqrt{3}a(\overline{X}-\overline{Y})}{\sqrt{S_1^2+2S_2^2}}\sim t(12).$$

对比系数则 $b=\sqrt{3}a=\pm\sqrt{\dfrac{135}{14}}$.

B11. 在两个等方差的正态总体中, 独立地各抽取一个样本容量为 7 的样本, 它们的样本方差分别为 S_1^2, S_2^2, 若 $P\left\{\max\left\{\dfrac{S_1^2}{S_2^2},\dfrac{S_2^2}{S_1^2}\right\}>c\right\}=0.05$, 求 c.

解: 设正态总体方差为 σ^2, 因为 $\dfrac{6S_1^2}{\sigma^2}\sim\chi^2(6)$, $\dfrac{6S_2^2}{\sigma^2}\sim\chi^2(6)$.则

$$F_1=\frac{S_1^2}{S_2^2}=\frac{\dfrac{6S_1^2}{\sigma^2}}{\dfrac{6S_2^2}{\sigma^2}}\sim F(6,6),\qquad F_2=\frac{S_2^2}{S_1^2}=\frac{\dfrac{6S_2^2}{\sigma^2}}{\dfrac{6S_1^2}{\sigma^2}}\sim F(6,6).$$

又因为 $F_2=\dfrac{1}{F_1}$, 则有

$$\begin{aligned}P\left\{\max\left\{\frac{S_1^2}{S_2^2},\frac{S_2^2}{S_1^2}\right\}>c\right\}&=P\{\max\{F_1,F_2\}>c\}=1-P\left\{\max\left\{F_1,F_2\right\}\leqslant c\right\}\\&=1-P\{F_1\leqslant c,F_2\leqslant c\}=1-P\left\{F_1\leqslant c,\frac{1}{F_1}\leqslant c\right\}\\&=1-P\left\{\frac{1}{c}\leqslant F_1\leqslant c\right\}=1-\left(P\left\{F_1\geqslant\frac{1}{c}\right\}-P\{F_1>c\}\right)\\&=1-P\left\{\frac{1}{F_1}\leqslant c\right\}+P\{F_1>c\}=1-(1-P\{F_2>c\})+P\{F_1>c\}\\&=P\{F_1>c\}+P\{F_2>c\}=0.05.\end{aligned}$$

又因为 F_1 与 F_2 同分布, 则 $P\{F_1>c\}=P\{F_2>c\}=0.025$.

此时查表得 $c=F_{0.025}(6,6)=5.82$.

B12. 设总体 $X\sim\chi^2(n)$, $X_1,X_2,\cdots,X_{16}$ 是来自总体 X 的简单随机样本.

(1) 求 $P\left\{\dfrac{\sum\limits_{i=1}^{8}X_i}{\sum\limits_{i=9}^{16}X_i}\leqslant 1\right\}$;

(2) 求 $P\left\{\dfrac{\sum\limits_{i=1}^{8}X_i}{\sum\limits_{i=9}^{16}X_i}= 1\right\}$.

解: (1) 由题, $Y_1=\sum\limits_{i=1}^{8}X_i\sim\chi^2(8n)$, $Y_2=\sum\limits_{i=9}^{16}X_i\sim\chi^2(8n)$.

从而 $\dfrac{Y_1}{Y_2}=\dfrac{\frac{Y_1}{8n}}{\frac{Y_2}{8n}}\sim F(8n,8n)$, 且此时 $\dfrac{Y_2}{Y_1}\sim F(8n,8n)$.

因为 $P\left\{\dfrac{Y_1}{Y_2}\leqslant 1\right\}=1-P\left\{\dfrac{Y_1}{Y_2}>1\right\}=1-P\left\{\dfrac{Y_2}{Y_1}<1\right\}$.

且由同分布可知 $P\left\{\dfrac{Y_1}{Y_2}\leqslant 1\right\}=P\left\{\dfrac{Y_2}{Y_1}\leqslant 1\right\}=P\left\{\dfrac{Y_2}{Y_1}<1\right\}$.

从而 $P\left\{\dfrac{Y_1}{Y_2}\leqslant 1\right\}=1-P\left\{\dfrac{Y_1}{Y_2}\leqslant 1\right\}$, 则 $P\left\{\dfrac{Y_1}{Y_2}\leqslant 1\right\}=0.5$, 即 $P\left\{\dfrac{\sum\limits_{i=1}^{8}X_i}{\sum\limits_{i=9}^{16}X_i}\leqslant 1\right\}=0.5$.

(2) 因为 $\dfrac{Y_1}{Y_2}$ 为连续性随机变量, 从而 $P\left\{\dfrac{\sum\limits_{i=1}^{8}X_i}{\sum\limits_{i=9}^{16}X_i}= 1\right\}=P\left\{\dfrac{Y_1}{Y_2}=1\right\}=0$.

第 7 章　参数估计

A1. 设 $X_1, X_2, \cdots, X_n$ 是来自下列总体 X 的简单随机样本, 求各总体分布中参数的矩估计量:

(1) $X \sim 0-1(p)$;　(2) $X \sim P(\lambda)$;　(3) $X \sim U(a, 2)$.

解: (1) 由于 $\mu_1 = E(X) = p$, 则用 A_1 代替 μ_1, 则矩估计量 $\hat{p} = A_1 = \overline{X}$.

(2) 由于 $\mu_1 = E(X) = \lambda$, 则用 A_1 代替 μ_1, 则矩估计量 $\hat{\lambda} = A_1 = \overline{X}$.

(3) 由于 $\mu_1 = E(X) = \dfrac{a+2}{2}$, 解得 $a = 2\mu_1 - 2$, 则用 A_1 代替 μ_1, 则矩估计量 $\hat{a} = 2A_1 - 2 = 2\overline{X} - 2$.

A2. 设 $X_1, X_2, \cdots, X_n$ 是来自下列总体 X 的简单随机样本, 求各总体分布中参数的极大似然估计量:

(1) $X \sim 0-1(p)$; (2) $X \sim E(\lambda)$; (3) $X \sim U(1, b)$.

解:

A3. 设 $X_1, X_2, \cdots, X_n$ 是来自下列总体 X 的简单随机样本, 求各总体中未知参数 θ 的矩估计量和极大似然估计量, 并对所获得的样本值, 求参数 θ 的矩估计值和极大似然估计值:

$$(1)\quad f(x;\theta) = \begin{cases} 2^{-\theta}\theta x^{\theta-1}, & 0 < x < 2, \\ 0, & 其他, \end{cases} \qquad \theta > 0;$$

样本值: 0.45 0.2 0.5 0.47 0.35 1.63 0.14 0.06 0.89 0.34

$$(2)\quad f(x;\theta) = \frac{1}{2\theta}\mathrm{e}^{-\frac{|x|}{\theta}}, -\infty < x < +\infty, \theta > 0;$$

样本值: −0.05 −0.47 0.01 −0.03 −0.18 1.65 −0.64 −1.05 0.41 −0.19

$$(3)\quad f(x;\theta) = \begin{cases} \dfrac{1}{2-\theta}, & \theta \leqslant x < 2, \\ 0, & 其他, \end{cases} \qquad \theta < 2;$$

样本值: 0.95 0.63 1.69 1.97 0.84 1.81 0.53 0.35 1.34 0.82

解: 设 $x_1, x_2, \cdots, x_n$ 是样本 $X_1, X_2, \cdots, X_n$ 的一组观测值.

(1) 由题

$$\mu_1 = E(X) = \int_0^2 x \cdot 2^{-\theta}\theta x^{\theta-1}\mathrm{d}x = 2^{-\theta}\theta \cdot \left.\frac{x^{\theta+1}}{\theta+1}\right|_0^2 = 2^{-\theta}\theta \cdot \frac{2^{\theta+1}}{\theta+1} = \frac{2\theta}{\theta+1}.$$

因此解得 $\theta=\dfrac{\mu_1}{2-\mu_1}$. 用 A_1 代替 μ_1, 则 θ 的矩估计量

$$\hat{\theta}_1=\frac{A_1}{2-A_1}=\frac{\overline{X}}{2-\overline{X}}.$$

由题可得似然函数 $L(\theta)=\prod\limits_{i=1}^{n}f(x_i;\theta)=\begin{cases}2^{-n\theta}\theta^n\prod\limits_{i=1}^{n}x_i^{\theta-1}, & 0<x_i<2,\\ 0, & \text{其他}.\end{cases}$

则对数似然函数: $l(\theta)=\ln L(\theta)=\begin{cases}-n\theta\ln 2+n\ln\theta+(\theta-1)\sum\limits_{i=1}^{n}\ln x_i, & 0<x_i<2,\\ 0, & \text{其他}.\end{cases}$

从而 $0<x_i<2$ 时, $\dfrac{\mathrm{d}l(\theta)}{\mathrm{d}\theta}=-n\ln 2+\dfrac{n}{\theta}+\sum\limits_{i=1}^{n}\ln x_i$.

令 $\left.\dfrac{\mathrm{d}l(\theta)}{\mathrm{d}\theta}\right|_{\theta=\hat{\theta}}=0$, 则 $-n\ln 2+\dfrac{n}{\hat{\theta}}+\sum\limits_{i=1}^{n}\ln x_i=0$. 解得 $\hat{\theta}=\dfrac{1}{\ln 2-\dfrac{1}{n}\sum\limits_{i=1}^{n}\ln x_i}$.

因此 θ 的极大似然估计量

$$\hat{\theta}_2=\frac{1}{\ln 2-\dfrac{1}{n}\sum\limits_{i=1}^{n}\ln X_i}.$$

对本题样本值, 有 $\overline{x}=\dfrac{1}{10}\sum\limits_{i=1}^{10}x_i=0.503$, 从而矩估计值为 $\hat{\theta}_1=\dfrac{\overline{x}}{2-\overline{x}}=\dfrac{0.503}{2-0.503}=0.336$.

又因为此时 $\sum\limits_{i=1}^{10}\ln x_i=-10.392$,

从而极大似然估计值为 $\hat{\theta}_2=\dfrac{1}{\ln 2-\dfrac{1}{10}\sum\limits_{i=1}^{10}\ln x_i}=\dfrac{1}{0.6931-(-1.0392)}=0.577$.

(2) 由题, 由奇函数的性质

$$\mu_1=E(X)=\int_{-\infty}^{+\infty}x\cdot\frac{1}{2\theta}e^{-\frac{|x|}{\theta}}=0.$$

此时 μ_1 与 θ 无关. 故应该继续考虑二阶矩, 有

$$\begin{aligned}\mu_2=E(X^2)&=\int_{-\infty}^{+\infty}x^2\cdot\frac{1}{2\theta}\mathrm{e}^{-\frac{|x|}{\theta}}\,\mathrm{d}x=\int_{0}^{+\infty}x^2e^{-\frac{x}{\theta}}\,\mathrm{d}\frac{x}{\theta}\xlongequal{u=\frac{x}{\theta}}\theta^2\int_0^{+\infty}u^2e^{-u}\mathrm{d}u\\&=\theta^2\Gamma(3)=2\theta^2.\qquad\text{(这里利用了 }\Gamma\text{ 函数的定义, 下同)}\end{aligned}$$

因为 $\theta>0$, 从而有 $\theta=\sqrt{\dfrac{\mu_2}{2}}$. 用 A_2 代替 μ_2, 则 θ 的矩估计量

$$\hat{\theta}_1=\sqrt{\frac{A_2}{2}}=\sqrt{\frac{1}{2n}\sum_{i=1}^{n}X_i^2}.$$

由题可得似然函数 $L(\theta)=\prod_{i=1}^{n}f(x_i;\theta)=\frac{1}{2^n\theta^n}\prod_{i=1}^{n}\mathrm{e}^{-\frac{|x_i|}{\theta}}$.

则对数似然函数 $l(\theta)=\ln L(\theta)=-n\ln 2-n\ln\theta-\frac{1}{\theta}\sum_{i=1}^{n}|x_i|$.

从而 $\frac{\mathrm{d}l(\theta)}{\mathrm{d}\theta}=-\frac{n}{\theta}+\frac{1}{\theta^2}\sum_{i=1}^{n}|x_i|$, 从而 $\left.\frac{\mathrm{d}l(\theta)}{\mathrm{d}\theta}\right|_{\theta=\hat{\theta}}=0$ 时有 $\hat{\theta}=\frac{1}{n}\sum_{i=1}^{n}|x_i|$.

因此 θ 的极大似然估计量

$$\hat{\theta_2}=\frac{1}{n}\sum_{i=1}^{n}|X_i|.$$

对本题样本值, $\sum_{i=1}^{10}x_i^2=4.6956$, $\sum_{i=1}^{10}|x_i|=4.68$,

因此矩估计值 $\hat{\theta_1}=\sqrt{\frac{1}{20}\sum_{i=1}^{10}x_i^2}=\sqrt{\frac{4.6956}{20}}=0.4845$. 极大似然估计值 $\hat{\theta_2}=\frac{1}{10}\sum_{i=1}^{10}|x_i|=0.468$.

(3) 由题

$$\mu_1=E(X)=\int_{\theta}^{2}\frac{x}{2-\theta}\mathrm{d}x=\left.\frac{x^2}{2(2-\theta)}\right|_{\theta}^{2}=\frac{2^2-\theta^2}{2(2-\theta)}=\frac{\theta+2}{2}.$$

解得 $\theta=2\mu_1-2$, 用 A_1 代替 μ_1, 则 θ 的矩估计量

$$\hat{\theta_1}=2A_1-2=2\overline{X}-2.$$

由题可得似然函数 $L(\theta)=\prod_{i=1}^{n}f(x_i;\theta)=\begin{cases}\frac{1}{(2-\theta)^n}, & \theta\leqslant x_i<2,\\ 0, & \text{其他}.\end{cases}$

则对数似然函数 $l(\theta)=\ln L(\theta)=\begin{cases}-n\ln(2-\theta), & \theta\leqslant x_i<2,\\ 0, & \text{其他}.\end{cases}$. 从而 $\theta\leqslant x_i<2$ 时 $\frac{\mathrm{d}l(\theta)}{\mathrm{d}\theta}=\frac{n}{2-\theta}>0$.

即对数似然函数递增, 从而当 $l(\theta)$ 最大时, θ 要取到最大.

由题, $\theta\leqslant x_i$ 成立, 从而 $\theta\leqslant\min\{x_1,x_2,\cdots,x_n\}$. 则此时应取 $\theta=\min\{x_1,x_2,\cdots,x_n\}$.

因此 θ 的极大似然估计量

$$\hat{\theta_2}=\min\{X_1,X_2,\cdots,X_n\}.$$

对本题样本值, $\overline{x}=\frac{1}{10}\sum_{i=1}^{10}x_i=1.093$, 从而矩估计值为 $\hat{\theta_1}=2\overline{x}-2=0.186$.

又因为 $\min\{x_1,x_2,\cdots,x_{10}\}=0.35$, 从而极大似然估计值为 $\hat{\theta_2}=0.35$.

A4. 设总体 X 服从参数为 p 的几何分布, 具有概率分布律

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

$X_1,X_2,\cdots,X_n$ 是来自该总体的一个样本, 给定的样本观测值为 $x_1,x_2,\cdots,x_n$, 求参数 p 的极大似然估计值.

解: 由题可知, 似然函数

$$L(p)=\prod_{i=1}^{n}p(x_i;p)=p^n\prod_{i=1}^{n}(1-p)^{x_i-1}.$$

因此对数似然函数 $l(p)=\ln L(p)=n\ln p+\sum_{i=1}^{n}(x_i-1)\ln(1-p)=n\ln p-n\ln(1-p)+\sum_{i=1}^{n}x_i\ln(1-p)$.

从而 $\dfrac{\mathrm{d}l(p)}{\mathrm{d}p}=\dfrac{n}{p}+\dfrac{n}{1-p}-\dfrac{1}{1-p}\sum_{i=1}^{n}x_i=\dfrac{n}{p(1-p)}-\dfrac{1}{1-p}\sum_{i=1}^{n}x_i$.

令 $\left.\dfrac{\mathrm{d}l(p)}{\mathrm{d}p}\right|_{p=\hat{p}}=0$, 则 $\dfrac{n}{\hat{p}(1-\hat{p})}-\dfrac{1}{1-\hat{p}}\sum_{i=1}^{n}x_i=0$. 解得 $\hat{p}=\dfrac{n}{\sum_{i=1}^{n}x_i}$.

因此 p 的极大似然估计值即为 $\hat{p}=\dfrac{n}{\sum_{i=1}^{n}x_i}$.

A5. 设总体 X 具有如下概率分布律:

X	0	1	2
p	θ^2	$2\theta(1-\theta)$	$(1-\theta)^2$

其中 $0<\theta<1$. 从上述总体中抽取样本容量为 9 的简单随机样本, 观测值: 2, 0, 2, 1, 0, 0, 1, 2, 1, 求参数 θ 的矩估计值和极大似然估计值.

解: 由题

$$\mu_1=E(X)=0\cdot\theta^2+1\cdot 2\theta(1-\theta)+2(1-\theta)^2=2-2\theta.$$

解得 $\theta=1-\dfrac{\mu_1}{2}$. 由样本得 $A_1=\dfrac{1}{9}\sum_{i=1}^{9}x_i=1$, 用 A_1 代替 μ_1.

则 θ 的矩估计值 $\hat{\theta_1}=1-\dfrac{A_1}{2}=1-\dfrac{1}{2}=0.5$.

由样本的观测值可知, 0,1,2 都各出现三次, 则似然函数

$$L(\theta)=\prod_{i=1}^{9}p(x_i;\theta)=(\theta^2)^3\cdot[2\theta(1-\theta)]^3\cdot[(1-\theta)^2]^3=8\theta^9(1-\theta)^9.$$

因此对数似然函数 $l(\theta)=\ln L(\theta)=\ln 8+9\ln\theta+9\ln(1-\theta)$.

从而 $\dfrac{\mathrm{d}l(\theta)}{\mathrm{d}\theta}=\dfrac{9}{\theta}-\dfrac{9}{1-\theta}$, 当 $\left.\dfrac{\mathrm{d}l(\theta)}{\mathrm{d}\theta}\right|_{\theta=\hat{\theta}}=0$ 时有 $\hat{\theta}=1-\hat{\theta}$, 解得 $\hat{\theta}=0.5$.

因此 θ 的极大似然估计值为 $\hat{\theta}_2=0.5$.

A6. 设总体 X 具有如下概率分布律:

X	0	1	2
p	θ	λ	$1-\theta-\lambda$

其中 $0<\theta<1, 0<\lambda<1, 0<\theta+\lambda<1$. 从上述总体中抽取样本容量为 9 的简单随机样本, 观察值: 2, 0, 2, 1, 0, 0, 1, 2, 1, 求参数 θ 和 λ 的矩估计值和极大似然估计值.

解: 由题

$$\mu_1 = E(X) = 0\cdot\theta + 1\cdot\lambda + 2(1-\theta-\lambda) = 2-2\theta-\lambda.$$

$$\mu_2 = E(X^2) = 0\cdot\theta + 1^2\cdot\lambda + 2^2(1-\theta-\lambda) = 4-4\theta-3\lambda.$$

联立解得 $\theta = 1-\dfrac{3}{2}\mu_1+\dfrac{1}{2}\mu_2,\ \lambda = 2\mu_1-\mu_2.$

由样本得 $A_1 = \dfrac{1}{9}\sum\limits_{i=1}^{9}x_i = 1,\ A_2 = \dfrac{1}{9}\sum\limits_{i=1}^{9}x_i^2 = \dfrac{15}{9} = \dfrac{5}{3}$. 用 A_1 代替 μ_1, A_2 代替 μ_2.

则 θ 的矩估计值 $\hat{\theta_1} = 1-\dfrac{3}{2}A_1+\dfrac{1}{2}A_2 = 1-\dfrac{3}{2}+\dfrac{5}{6} = \dfrac{1}{3}$.

λ 的矩估计值 $\hat{\lambda_1} = 2A_1-A_2 = \dfrac{1}{3}$.

由样本的观测值可知, 0,1,2 都各出现三次, 则似然函数

$$L(\theta,\lambda) = \prod_{i=1}^{9}p(x_i;\theta,\lambda) = \theta^3\cdot\lambda^3\cdot(1-\theta-\lambda)^3.$$

因此对数似然函数 $l(\theta,\lambda) = \ln L(\theta,\lambda) = 3\ln\theta+3\ln\lambda+3\ln(1-\theta-\lambda)$.

从而 $\dfrac{\partial l(\theta,\lambda)}{\partial\theta} = \dfrac{3}{\theta}-\dfrac{3}{1-\theta-\lambda},\ \dfrac{\partial l(\theta,\lambda)}{\partial\lambda} = \dfrac{3}{\lambda}-\dfrac{3}{1-\theta-\lambda}$.

当 $\left.\dfrac{\partial l(\theta,\lambda)}{\partial\theta}\right|_{\theta=\hat{\theta},\lambda=\hat{\lambda}} = \left.\dfrac{\partial l(\theta,\lambda)}{\partial\lambda}\right|_{\theta=\hat{\theta},\lambda=\hat{\lambda}} = 0$ 时有 $\hat{\theta} = \hat{\lambda} = 1-\hat{\theta}-\hat{\lambda}$.

解得 $\hat{\theta} = \hat{\lambda} = \dfrac{1}{3}$.

因此 θ 的极大似然估计值为 $\hat{\theta_2} = \dfrac{1}{3}$; λ 的极大似然估计值为 $\hat{\lambda_2} = \dfrac{1}{3}$.

A7. 设总体 X 的密度函数为

$$f(x;\theta) = \begin{cases} \dfrac{x}{\theta}\mathrm{e}^{-\frac{x^2}{2\theta}}, & x>0, \\ 0, & \text{其他}, \end{cases}$$

$\theta>0$ 未知, 记 $\mu_2 = E(X^2)$, $p = P\{X>1\}$. $X_1, X_2, \cdots, X_n$ 是来自该总体的简单随机样本, 求参数 θ, μ_2 和 p 的极大似然估计量.

解: 设 $x_1,x_2,\cdots,x_n$ 是样本 $X_1,X_2,\cdots,X_n$ 的一组观测值.

从而似然函数 $L(\theta)=\prod_{i=1}^{n}f(x_i;\theta)=\begin{cases}\dfrac{1}{\theta^n}\prod_{i=1}^{n}x_i\mathrm{e}^{-\frac{x_i^2}{2\theta}}, & x_i>0,\\ 0, & \text{其他}.\end{cases}$

则对数似然函数 $l(\theta)=\ln L(\theta)=\begin{cases}-n\ln\theta+\sum_{i=1}^{n}\ln x_i-\dfrac{1}{2\theta}\sum_{i=1}^{n}x_i^2, & x_i>0,\\ 0, & \text{其他}.\end{cases}$

从而 $x_i>0$ 时, $\dfrac{\mathrm{d}l(\theta)}{\mathrm{d}\theta}=-\dfrac{n}{\theta}+\dfrac{1}{2\theta^2}\sum_{i=1}^{n}x_i^2$.

令 $\left.\dfrac{\mathrm{d}l(\theta)}{\mathrm{d}\theta}\right|_{\theta=\hat\theta}=0$, 则 $-\dfrac{n}{\hat\theta}+\dfrac{1}{2\hat\theta^2}\sum_{i=1}^{n}x_i^2=0$, 解得 $\hat\theta=\dfrac{1}{2n}\sum_{i=1}^{n}x_i^2$.

因此 θ 的极大似然估计量为 $\hat\theta=\dfrac{1}{2n}\sum_{i=1}^{n}X_i^2$.

又因为

$$\begin{aligned}\mu=E(X^2)&=\int_0^{+\infty}x^2\cdot\frac{x}{\theta}\mathrm{e}^{-\frac{x^2}{2\theta}}\mathrm{d}x=\int_0^{+\infty}x^2\cdot\mathrm{e}^{-\frac{x^2}{2\theta}}\mathrm{d}\frac{x^2}{2\theta}\overset{u=\frac{x^2}{2\theta}}{=\!=\!=}\int_0^{+\infty}2\theta u e^{-u}\mathrm{d}u\\&=2\theta\Gamma(2)=2\theta.\end{aligned}$$

因此 μ 的极大似然估计量 $\hat\mu=2\hat\theta=\dfrac{1}{n}\sum_{i=1}^{n}X_i^2$.

又因为

$$\begin{aligned}p=P\{X>1\}&=\int_1^{+\infty}f(x)\mathrm{d}x=\int_1^{+\infty}\frac{x}{\theta}e^{-\frac{x^2}{2\theta}}\mathrm{d}x=\int_1^{+\infty}e^{-\frac{x^2}{2\theta}}\mathrm{d}\frac{x^2}{2\theta}\\&=-e^{-\frac{x^2}{2\theta}}\Big|_1^{+\infty}=0-\left(-e^{-\frac{1}{2\theta}}\right)=e^{-\frac{1}{2\theta}}.\end{aligned}$$

因此 p 的极大似然估计量 $\hat p=\exp\left(-\dfrac{1}{2\hat\theta}\right)=\exp\left(-\dfrac{n}{\sum_{i=1}^{n}X_i^2}\right)$.

A8. 设总体 X 的均值为 μ, 方差为 $\sigma^2,X_1,X_2,\cdots,X_{10}$ 为来自总体 X 的简单随机样本, 问 a 取何值时, $a\sum_{i=1}^{9}(X_{i+1}-X_i)^2$ 是 σ^2 的无偏估计量?

解: 由题, 对 $i=1,2,\cdots,10$, $E(X_i)=\mu$, $E(X_i^2)=\mathrm{Var}(X_i)+(E(X_i))^2=\mu^2+\sigma^2$. 从而

$$\begin{aligned}E\left(a\sum_{i=1}^{9}(X_{i+1}-X_i)^2\right)&=aE\left[\sum_{i=1}^{9}(X_{i+1}^2-2X_{i+1}X_i+X_i^2)\right]\\&=a\sum_{i=1}^{9}[E(X_{i+1}^2)-2E(X_{i+1}X_i)+E(X_i^2)]\end{aligned}$$

$$
\begin{aligned}
&= a\sum_{i=1}^{9}[E(X_{i+1}^2) - 2E(X_{i+1})E(X_i) + E(X_i^2)] \\
&= a\sum_{i=1}^{9}[(\sigma^2+\mu^2) - 2\mu^2 + (\sigma^2+\mu^2)] \\
&= a\sum_{i=1}^{9} 2\sigma^2 = 18a\sigma^2.
\end{aligned}
$$

由无偏估计量可知, $18a\sigma^2 = E(\sigma^2) = \sigma^2$, 则 $a = \dfrac{1}{18}$.

A9. 设总体 $X \sim N(\mu, \sigma^2)$, μ, σ^2 未知, X_1, X_2, X_3 是总体 X 的简单随机样本, 用

$$
\begin{aligned}
\widehat{\mu}_1 &= \frac{1}{2}X_1 + \frac{1}{4}X_2 + \frac{1}{4}X_3, \\
\widehat{\mu}_2 &= 2X_1 - 2X_2 + X_3, \\
\widehat{\mu}_3 &= \frac{1}{3}X_1 + \frac{1}{3}X_2 + \frac{1}{3}X_3
\end{aligned}
$$

估计参数 μ, 它们都是无偏估计量吗? 如果是, 哪个更有效?

解: 由题有

$$
E(\widehat{\mu}_1) = E\left(\frac{1}{2}X_1 + \frac{1}{4}X_2 + \frac{1}{4}X_3\right) = \frac{1}{2}E(X_1) + \frac{1}{4}E(X_2) + \frac{1}{4}E(X_3) = \left(\frac{1}{2} + \frac{1}{4} + \frac{1}{4}\right)\mu = \mu.
$$

$$
E(\widehat{\mu}_2) = E(2X_1 - 2X_2 + X_3) = 2E(X_1) - 2E(X_2) + E(X_3) = 2\mu - 2\mu + \mu = \mu.
$$

$$
E(\widehat{\mu}_3) = E\left(\frac{1}{3}X_1 + \frac{1}{3}X_2 + \frac{1}{3}X_3\right) = \frac{1}{3}E(X_1) + \frac{1}{3}E(X_2) + \frac{1}{3}E(X_3) = \left(\frac{1}{3} + \frac{1}{3} + \frac{1}{3}\right)\mu = \mu.
$$

期望均为 μ, 则它们都是无偏估计量.

考虑有效性, 要计算它们的方差.

$$
\operatorname{Var}(\widehat{\mu}_1) = \operatorname{Var}\left(\frac{1}{2}X_1\right) + \operatorname{Var}\left(\frac{1}{4}X_2\right) + \operatorname{Var}\left(\frac{1}{4}X_3\right) = \frac{1}{4}\sigma^2 + \frac{1}{16}\sigma^2 + \frac{1}{16}\sigma^2 = \frac{3}{8}\sigma^2.
$$

$$
\operatorname{Var}(\widehat{\mu}_2) = \operatorname{Var}(2X_1) + \operatorname{Var}(-2X_2) + \operatorname{Var}(X_3) = 4\sigma^2 + 4\sigma^2 + \sigma^2 = 9\sigma^2.
$$

$$
\operatorname{Var}(\widehat{\mu}_3) = \operatorname{Var}\left(\frac{1}{3}X_1\right) + \operatorname{Var}\left(\frac{1}{3}X_2\right) + \operatorname{Var}\left(\frac{1}{3}X_3\right) = \frac{1}{9}\sigma^2 + \frac{1}{9}\sigma^2 + \frac{1}{9}\sigma^2 = \frac{1}{3}\sigma^2.
$$

因此 $\operatorname{Var}(\widehat{\mu}_3)$ 最小, 则 $\widehat{\mu}_3$ 最有效.

A10. 设 $\widehat{\theta}_1$ 和 $\widehat{\theta}_2$ 都是 θ 的无偏估计量, 且 $\widehat{\theta}_1$ 和 $\widehat{\theta}_2$ 相互独立. 已知 $D(\widehat{\theta}_1)=\sigma_1^2$, $D(\widehat{\theta}_2)=\sigma_2^2$, 引入 θ 的一个新的无偏估计量 $\widehat{\theta}_3=\alpha\widehat{\theta}_1+(1-\alpha)\widehat{\theta}_2$, 试确定常数 α, 使 $D(\widehat{\theta}_3)$ 达到最小.

解: 由独立可知

$$D(\widehat{\theta}_3)=D(\alpha\widehat{\theta}_1+(1-\alpha)\widehat{\theta}_2)=\alpha^2D(\widehat{\theta}_1)+(1-\alpha)^2D(\widehat{\theta}_2)=\alpha^2\sigma_1^2+(1-\alpha)^2\sigma_2^2.$$

对 $f(\alpha)=\alpha^2\sigma_1^2+(1-\alpha)^2\sigma_2^2=(\sigma_1^2+\sigma_2^2)\alpha^2-2\sigma_2^2\alpha+\sigma_2^2$, 有 $f'(\alpha)=2(\sigma_1^2+\sigma_2^2)\alpha-2\sigma_2^2$.

当 $f'(\alpha)=0$ 时, $\alpha=\dfrac{\sigma_2^2}{\sigma_1^2+\sigma_2^2}$, 且 $\alpha>\dfrac{\sigma_2^2}{\sigma_1^2+\sigma_2^2}$ 时 $f'(\alpha)>0$; $\alpha<\dfrac{\sigma_2^2}{\sigma_1^2+\sigma_2^2}$ 时 $f'(\alpha)<0$.

从而 $\alpha=\dfrac{\sigma_2^2}{\sigma_1^2+\sigma_2^2}$ 时 $f(\alpha)$ 取最小值, 即 $D(\widehat{\theta}_3)$ 最小.

A11. 某机器生产的螺杆直径 X (单位: mm) 服从正态分布 $N(\mu,0.3^2)$.

(1) 从总体中抽取样本容量为 5 的样本, 测得直径: 22.3, 21.5, 21.8, 21.4, 22.1, 试在 95% 的置信水平下求该机器所生产的螺杆平均直径 μ 的置信区间;

(2) 若要使螺杆的平均直径 μ 的置信水平为 95% 的置信区间长度不超过 0.2, 问样本容量 n 至少应取多大?

解:

A12. 某厂生产的灯泡寿命 X (单位: h) 服从正态分布 $N(\mu,\sigma^2)$, μ,σ^2 未知, 从已生产的一大批灯泡中采用无放回抽样方式抽取 15 只, 测得其寿命如下:

4040 2990 2964 3245 3026 3633 3387 4136 3595 3194 3714 2831 3845 3410 3004

(1) 求 μ 的置信水平为 95% 的置信区间;

(2) 求 μ 的置信水平为 95% 的单侧置信下限.

解:

A13. 某医学研究所研发了一种新药, 对 9 个试验者进行为期半年的观察, 记录服用该药前后的甘油三酯水平的变化, 数据如下 (单位: mg/dL):

服药前	180	139	152	167	138	160	107	156	94
服药后	100	92	118	171	132	123	84	112	105

假设服药前后甘油三酯水平差服从正态分布, 在置信水平 95% 下给出服药前后甘油三酯平均水平差的区

间估计.

解:

A14. 为比较甲、乙两种肥料对产量的影响, 研究者选择了 10 块田地, 将每块田地分成大小相同的两块, 随机选择一块用甲肥料, 另一块用乙肥料, 其他条件保持相同, 得到的产量 (单位: kg) 数据如下:

甲肥料	109	98	97	100	104	102	94	99	103	108
乙肥料	107	105	110	118	109	113	111	95	112	101

假设甲、乙两种肥料的产量差服从正态分布, 试在 95% 的置信水平下推断甲、乙两种肥料的平均产量差值的范围.

解:

A15. 下面 16 个数字来自计算机的正态随机数生成器:

8.801	3.817	8.223	6.374	9.252	7.352	13.781	7.599
13.134	4.465	6.533	7.021	9.015	7.325	7.041	9.560

(1) 给出总体均值 μ 和方差 σ^2 的极大似然估计值;

(2) 求均值 μ 的置信水平为 95% 的置信区间;

(3) 求方差 σ^2 的置信水平为 95% 的置信区间.

解:

A16. 已知某种电子管使用寿命 (单位: h) 服从正态分布 $N(\mu,\sigma^2)$, μ,σ^2 未知, 从一批电子管中随机抽取 16 只, 检测结果得样本标准差为 300 h. 在置信水平 95% 下求:

(1) σ 的置信区间;

(2) σ 的单侧置信上限.

解:

A17. 为了解某市两所高校学生的消费情况, 在两所高校各随机调查 100 人, 调查结果: 甲校学生月平均消费 803 元, 标准差 75 元; 乙校学生月平均消费 938 元, 标准差 102 元. 假设甲校学生月平均消费额 (单位: 元) $X\sim N(\mu_1,\sigma^2)$, 乙校学生月平均消费额 (单位: 元) $Y\sim N(\mu_2,\sigma^2)$, μ_1,μ_2,σ^2 未知, 两样本相互独立, 求两校学生月平均消费额差值 $\mu_1-\mu_2$ 的置信水平为 95% 的置信区间和单侧置信下限.

解:

A18. 某厂的一台瓶装灌装机, 每瓶的净重 X 服从正态分布 $N(\mu_1, \sigma_1^2)$, 从中随机抽出 16 瓶, 称得其净重的平均值为 456.64 g, 标准差为 12.8 g; 现引进一台新灌装机, 其每瓶的净重 Y 服从正态分布 $N(\mu_2, \sigma_2^2)$, 从中随机抽出 12 瓶, 称得其净重的平均值为 451.34 g, 标准差为 11.3 g.

(1) 假设 $\sigma_1 = 13, \sigma_2 = 12$, 求 $\mu_1 - \mu_2$ 的置信水平为 95% 的置信区间;

(2) 假设 $\sigma_1 = \sigma_2$ 未知, 求 $\mu_1 - \mu_2$ 的置信水平为 95% 的置信区间;

(3) 求 $\dfrac{\sigma_1^2}{\sigma_2^2}$ 的置信水平为 95% 的置信区间.

解:

A19. 某超市负责人需要比较郊区 A 和郊区 B 居民的平均收入以确定合适的分店地址. 假设两郊区居民的收入均服从正态分布, 对两郊区居民分别进行抽样调查, 各抽取 64 户家庭, 计算得郊区 A 居民的人均年收入为 3.276 万元, 标准差为 0.203 万元; 郊区 B 居民的人均年收入为 3.736 万元, 标准差为 0.421 万元. 假设两个正态总体的方差不相等, 求两郊区居民人均年收入平均差值的置信水平为 95% 的近似置信区间.

解:

B1. 设总体 X 的密度函数为

$$f(x;\theta)=\begin{cases}\dfrac{6x(\theta-x)}{\theta^3}, & 0<x<\theta,\\ 0, & \text{其他},\end{cases}$$

其中 $\theta>0$ 未知, $X_1,X_2,\cdots,X_n$ 是来自总体 X 的简单随机样本, 求 θ 的矩估计量 $\widehat{\theta}$, 并计算 $E(\widehat{\theta})$ 和 $\operatorname{Var}(\widehat{\theta})$.

解: 由题

$$\mu_1=E(X)=\int_0^\theta x\cdot\frac{6x(\theta-x)}{\theta^3}\mathrm{d}x=\left.\frac{2\theta x^3-\frac{3}{2}x^4}{\theta^3}\right|_0^\theta=\frac{2\theta^4-\frac{3}{2}\theta^4}{\theta^3}=\frac{\theta}{2}.$$

解得 $\theta=2\mu_1$, 用 A_1 代替 μ_1, 则 θ 的矩估计量

$$\hat{\theta}=2A_1=2\overline{X}.$$

因为 $E(\overline{X})=E(X)=\dfrac{\theta}{2}$, 从而 $E(\hat{\theta})=E(2\overline{X})=\theta$.

又因为

$$E(X^2)=\int_0^\theta x^2\cdot\frac{6x(\theta-x)}{\theta^3}\mathrm{d}x=\int_0^\theta\frac{-6x^4+6\theta x^3}{\theta^3}\mathrm{d}x=\left.\frac{-\frac{6}{5}x^5+\frac{3\theta}{2}x^4}{\theta^3}\right|_0^\theta=\frac{-\frac{6}{5}\theta^5+\frac{3}{2}\theta^5}{\theta^3}=\frac{3\theta^2}{10}.$$

从而 $\operatorname{Var}(X)=E(X^2)-(E(X))^2=\dfrac{3\theta^2}{10}-\dfrac{\theta^2}{4}=\dfrac{\theta^2}{20}$.

因此

$$\operatorname{Var}(\hat{\theta})=\operatorname{Var}(2\overline{X})=4\operatorname{Var}\left(\frac{1}{n}\sum_{i=1}^n X_i\right)=\frac{4}{n^2}\sum_{i=1}^n\operatorname{Var}(X_i)=\frac{4}{n^2}\cdot n\cdot\frac{\theta^2}{20}=\frac{\theta^2}{5n}.$$

B2. 设湖中有 N 条鱼 (N 未知), 现钓出 r 条, 做上记号后放回湖中, 一段时间后, 再钓出 S 条 ($S\geqslant r$), 结果发现其中 t 条有记号, 试用极大似然法估计湖中鱼的数量 N.

解: 第一次钓完做记号后, 湖中每条鱼带有记号的概率 $p=\dfrac{r}{N}$.

记第二次钓出的鱼中有记号的数量为 X. 则此时应有 $X\sim B(S,p)$.

有 $P\{X=t\}=\mathrm{C}_S^t p^t(1-p)^{S-t}=\mathrm{C}_S^t\dfrac{r^t}{N^t}\cdot\dfrac{(N-r)^{S-t}}{N^{S-t}}=\mathrm{C}_S^t\dfrac{r^t(N-r)^{S-t}}{N^S}$.

此时相当于只进行了一次抽样实验, 则似然函数 $L(N)=P\{X=t\}$.

对数似然函数 $l(N)=\ln\mathrm{C}_S^t+t\ln r+(S-t)\ln(N-r)-S\ln N$. 求导有 $\dfrac{\mathrm{d}l(N)}{\mathrm{d}N}=\dfrac{S-t}{N-r}-\dfrac{S}{N}$.

令 $\left.\dfrac{\mathrm{d}l(N)}{\mathrm{d}N}\right|_{N=\hat{N}}=0$. 此时有 $\dfrac{S-t}{\hat{N}-r}-\dfrac{S}{\hat{N}}=0$, 即 $\hat{N}(S-t)=S(\hat{N}-r)$, 解得 $\hat{N}=\dfrac{Sr}{t}$.

即 N 的极大似然估计量为 $\hat{N}=\dfrac{Sr}{t}$. 又结合实际情况, 鱼的数量应为整数.

所以对该估计值进行取整. 则湖中鱼数量 N 的极大似然法估计值为 $\left[\frac{Sr}{t}\right]$.

B3. 设总体 X 具有如下概率分布律:

X	a_1	a_2	a_3
p	θ	$\frac{1-\theta}{2}$	$\frac{1-\theta}{2}$

从总体 X 中取得样本容量为 n 的样本 $X_1, X_2, \cdots, X_n$, 记其中取 a_1, a_2, a_3 的个数分别为 n_1, n_2, n_3, 其中 $n_1+n_2+n_3=n$. 求参数 θ 的矩估计量和极大似然估计量.

解: 由题

$$\mu_1 = E(X) = a_1\theta + a_2 \cdot \frac{1-\theta}{2} + a_3 \cdot \frac{1-\theta}{2} = \frac{(a_2+a_3)+(2a_1-a_2-a_3)\theta}{2}.$$

解得 $\theta = \dfrac{2\mu_1 - a_2 - a_3}{2a_1 - a_2 - a_3}$, 用 A_1 代替 μ_1, 则 θ 的矩估计量

$$\widehat{\theta}_1 = \frac{2A_1 - a_2 - a_3}{2a_1 - a_2 - a_3} = \frac{2\overline{X} - a_2 - a_3}{2a_1 - a_2 - a_3}.$$

由题可知, 似然函数

$$\begin{aligned} L(\theta) &= \prod_{i=1}^{n} p(x_i;\theta) = (P\{X=a_1\})^{n_1} \cdot (P\{X=a_2\})^{n_2} \cdot (P\{X=a_3\})^{n_3} \\ &= \theta^{n_1}\left(\frac{1-\theta}{2}\right)^{n_2}\left(\frac{1-\theta}{2}\right)^{n_3} = \theta^{n_1}\left(\frac{1-\theta}{2}\right)^{n-n_1}. \end{aligned}$$

因此对数似然函数 $l(\theta) = \ln L(\theta) = n_1 \ln\theta + (n-n_1)\ln\dfrac{1-\theta}{2} = n_1\ln\theta + (n-n_1)\ln(1-\theta) - (n-n_1)\ln 2$.

从而 $\dfrac{\mathrm{d}l(\theta)}{\mathrm{d}\theta} = \dfrac{n_1}{\theta} - \dfrac{n-n_1}{1-\theta} = \dfrac{n_1(1-\theta)-(n-n_1)\theta}{\theta(1-\theta)} = \dfrac{n_1 - n\theta}{\theta(1-\theta)}$.

令 $\left.\dfrac{\mathrm{d}l(\theta)}{\mathrm{d}\theta}\right|_{\theta=\hat{\theta}} = 0$, 则 $n_1 - n\hat{\theta} = 0$. 解得 $\hat{\theta} = \dfrac{n_1}{n}$.

则 θ 的极大似然估计量

$$\widehat{\theta}_2 = \frac{n_1}{n}.$$

B4. 设总体 $X \sim N(\mu, \sigma^2)$, μ, σ^2 未知, 抽取三个独立样本 (X_1, X_2), (Y_1, Y_2, Y_3), (Z_1, Z_2, Z_3, Z_4), S_1^2, S_2^2, S_3^2 分别是对应的样本方差. 设 $T = aS_1^2 + bS_2^2 + cS_3^2$, 其中 a, b, c 是在区间 $[0,1]$ 上取值的实数.

(1) 写出 T 是 σ^2 的无偏估计量的充要条件;

(2) 问 a, b, c 取何值时, T 为最有效估计量?

解: (1) 由于 $E(S_1^2) = E(S_2^2) = E(S_3^2) = \sigma^2$, 则 $E(T) = E(aS_1^2 + bS_2^2 + cS_3^2) = (a+b+c)\sigma^2$.

由无偏估计量知, $E(T) = \sigma^2$. 当 $(a+b+c)\sigma^2 = \sigma^2$ 时有 $a+b+c=1$.

因此 T 是 σ^2 的无偏估计量的充要条件为 $a+b+c=1$.

(2) T 最有效时, $\mathrm{Var}(T)$ 最小.

因为 $X\sim N(\mu,\sigma^2)$, 则 $\dfrac{S_1^2}{\sigma^2}\sim\chi^2(1)$, $\dfrac{2S_2^2}{\sigma^2}\sim\chi^2(2)$, $\dfrac{3S_3^2}{\sigma^2}\sim\chi^2(3)$.

从而 $\mathrm{Var}\left(\dfrac{S_1^2}{\sigma^2}\right)=2$, $\mathrm{Var}\left(\dfrac{2S_2^2}{\sigma^2}\right)=4$, $\mathrm{Var}\left(\dfrac{3S_3^2}{\sigma^2}\right)=6$, 则有 $\mathrm{Var}(S_1^2)=2\sigma^4$, $\mathrm{Var}(S_2^2)=\sigma^4$, $\mathrm{Var}(S_3^2)=\dfrac{2\sigma^4}{3}$.

又因为 S_1^2, S_2^2, S_3^2 相互独立, 则 $\mathrm{Var}(T)=a^2\mathrm{Var}(S_1^2)+b^2\mathrm{Var}(S_2^2)+c^2\mathrm{Var}(S_3^2)=\dfrac{\sigma^4}{3}(6a^2+3b^2+2c^2)$.

设 $L(a,b,c,\lambda)=6a^2+3b^2+2c^2+\lambda(a+b+c-1)$.

则 $L'_a=12a+\lambda$, $L'_b=6b+\lambda$, $L'_c=4c+\lambda$. 当 $L'_a=L'_b=L'_c=0$ 时则 $12a=6b=4c$, 即 $b=2a$, $c=3a$.

由 $a+b+c=1$ 则 $a=\dfrac{1}{6}$, $b=\dfrac{1}{3}$, $c=\dfrac{1}{2}$, 此时 $6a^2+3b^2+2c^2$ 取最小值.

也即 $a=\dfrac{1}{6}$, $b=\dfrac{1}{3}$, $c=\dfrac{1}{2}$ 时 T 为最有效估计量.

B5. 设总体 X 的密度函数为

$$f(x;\theta)=\begin{cases}\dfrac{2x}{\theta^2}, & 0<x<\theta,\\ 0, & \text{其他},\end{cases}$$

$\theta>0$ 未知, $X_1,X_2,\ldots,X_n(n\geqslant 4)$ 是来自总体 X 的简单随机样本.

(1) 求 θ 的矩估计量 $\widehat{\theta}_1$, 和极大似然估计量 $\widehat{\theta}_2$;

(2) 在均方误差准则下, 判断哪个估计量更优;

(3) 判断两个估计量是否为 θ 的相合估计量.

解: (1) 由题

$$\mu_1=E(X)=\int_0^\theta x\cdot\frac{2x}{\theta^2}\mathrm{d}x=\left.\frac{2x^3}{3\theta^2}\right|_0^\theta=\frac{2\theta}{3}.$$

解得 $\theta=\dfrac{3\mu_1}{2}$, 用 A_1 代替 μ_1, 则 θ 的矩估计量

$$\widehat{\theta}_1=\frac{3A_1}{2}=\frac{3\overline{X}}{2}.$$

由题可得似然函数 $L(\theta)=\prod\limits_{i=1}^n f(x_i;\theta)=\begin{cases}\dfrac{2^n}{\theta^{2n}}\prod\limits_{i=1}^n x_i, & 0<x_i<\theta,\\ 0, & \text{其他}.\end{cases}$

则对数似然函数 $l(\theta)=\ln L(\theta)=\begin{cases}n\ln 2-2n\ln\theta+\sum\limits_{i=1}^n\ln x_i, & 0<x_i<\theta,\\ 0, & \text{其他}.\end{cases}.$

从而 $0<x_i<\theta$ 时 $\dfrac{\mathrm{d}l(\theta)}{\mathrm{d}\theta}=-\dfrac{2n}{\theta}<0$.

即对数似然函数递减, 从而当 $l(\theta)$ 最大时, θ 要取到最小.

由题, $\theta>x_i$ 成立, 从而 $\theta>\max\{x_1,x_2,\cdots,x_n\}$. 则此时应取 $\theta=\max\{x_1,x_2,\cdots,x_n\}$.

因此 θ 的极大似然估计量

$$\widehat{\theta}_2=\max\{X_1,X_2,\cdots,X_n\}.$$

(2) 分别计算二者的均方误差. 由题

$$E(X^2)=\int_0^\theta x^2\cdot\frac{2x}{\theta^2}\mathrm{d}x=\left.\frac{2x^4}{4\theta^2}\right|_0^\theta=\frac{\theta^2}{2}.$$

从而 $\operatorname{Var}(X)=E(X^2)-(E(X))^2=\dfrac{\theta^2}{2}-\dfrac{4\theta^2}{9}=\dfrac{\theta^2}{18}$.

又因为 $E(\widehat{\theta}_1)=\dfrac{3}{2}E(\overline{X})=\dfrac{3}{2}\cdot\dfrac{2\theta}{3}=\theta$, 则有

$$\operatorname{Mse}(\widehat{\theta}_1)=\operatorname{Var}(\widehat{\theta}_1)=\frac{9}{4}\operatorname{Var}(\overline{X})=\frac{9}{4}\cdot\frac{1}{n^2}\sum_{i=1}^n\operatorname{Var}(X_i)=\frac{9}{4n^2}\cdot n\cdot\frac{\theta^2}{18}=\frac{\theta^2}{8n}.$$

此时要考察 $M=\max\{X_1,X_2,\cdots,X_n\}$ 的分布.

因为 X 的分布函数 $F_X(x)=\begin{cases}1, & x\geqslant\theta,\\ \dfrac{x^2}{\theta^2}, & 0<x<\theta,\\ 0, & x\leqslant 0.\end{cases}$

从而 M 的分布函数 $F_M(x)=\displaystyle\prod_{i=1}^n F_{X_i}(x)=\begin{cases}1, & x\geqslant\theta,\\ \dfrac{x^{2n}}{\theta^{2n}}, & 0<x<\theta,\\ 0, & x\leqslant 0.\end{cases}$

则其密度函数 $f_M(x)=F_M'(x)=\begin{cases}\dfrac{2nx^{2n-1}}{\theta^{2n}}, & 0<x<\theta,\\ 0, & \text{其他}.\end{cases}$

此时有

$$E(M)=\int_0^\theta x\cdot\frac{2nx^{2n-1}}{\theta^{2n}}\mathrm{d}x=\left.\frac{2nx^{2n+1}}{(2n+1)\theta^{2n}}\right|_0^\theta=\frac{2n\theta}{2n+1}.$$

$$E(M^2)=\int_0^\theta x^2\cdot\frac{2nx^{2n-1}}{\theta^{2n}}\mathrm{d}x=\left.\frac{2nx^{2n+2}}{(2n+2)\theta^{2n}}\right|_0^\theta=\frac{n\theta^2}{n+1}.$$

从而

$$\begin{aligned}\operatorname{Mse}(\widehat{\theta}_2)&=E[(\widehat{\theta}_2-\theta)^2]=E(\widehat{\theta}_2^2)-2\theta E(\widehat{\theta}_2)+\theta^2=E(M^2)-2\theta E(M)+\theta^2\\&=\frac{n\theta^2}{n+1}-\frac{2\theta\cdot 2n\theta}{2n+1}+\theta^2=\frac{n(2n+1)-4n(n+1)+(n+1)(2n+1)}{(n+1)(2n+1)}\theta^2\\&=\frac{2n^2+n-4n^2-4n+2n^2+3n+1}{(n+1)(2n+1)}\theta^2=\frac{\theta^2}{(n+1)(2n+1)}.\end{aligned}$$

因为 $n \geqslant 4$, 则 $2n+1 \geqslant 9 > 8$, $n+1 > n$, 有 $(n+1)(2n+1) > 8n$, 从而 $\mathrm{Mse}(\widehat{\theta}_2) < \mathrm{Mse}(\widehat{\theta}_1)$.

从而在均方误差准则下, $\widehat{\theta}_2$ 优于 $\widehat{\theta}_1$.

(3) 由辛钦大数定律, $n \to +\infty$ 时, $\overline{X} \xrightarrow{P} E(X) = \dfrac{2}{3}\theta$.

因此 $\widehat{\theta}_1 = \dfrac{3}{2}\overline{X} \xrightarrow{P} \dfrac{3}{2} \cdot \dfrac{2}{3}\theta = \theta$.

即 $\widehat{\theta}_1$ 是 θ 的相合估计量.

由题有 $\widehat{\theta}_2 < \theta$, 此时有

$$E(|\widehat{\theta}_2 - \theta|) = E(\theta - \widehat{\theta}_2) = \theta - \frac{2n\theta}{2n+1} = \frac{\theta}{2n+1}.$$

从而由马尔可夫不等式, 对任意 $\varepsilon > 0$, 有

$$P\{|\widehat{\theta}_2 - \theta| < \varepsilon\} \geqslant 1 - \frac{E(|\widehat{\theta}_2 - \theta|)}{\varepsilon} = 1 - \frac{\theta}{(n+1)\varepsilon}.$$

因为 $n \to +\infty$ 时, $1 - \dfrac{\theta}{(n+1)\varepsilon} \to 1$, 且 $P\{|\widehat{\theta}_2 - \theta| < \varepsilon\} \leqslant 1$.

由夹逼准则, $\lim\limits_{n \to +\infty} P\{|\widehat{\theta}_2 - \theta| < \varepsilon\} = 1$, 从而 $\widehat{\theta}_2 \xrightarrow{P} \theta$.

即 $\widehat{\theta}_2$ 是 θ 的相合估计量.

B6. 设总体 X 的密度函数为

$$f(x;\theta) = \begin{cases} \dfrac{1}{\theta}\mathrm{e}^{-\frac{x}{\theta}}, & x > 0, \\ 0, & \text{其他}, \end{cases}$$

$\theta > 0$ 未知, $X_1, X_2, \cdots, X_n$ 是来自总体 X 的简单随机样本.

(1) 证明: 样本均值是 θ 的矩估计量, 也是极大似然估计量;

(2) 在形如 $c\sum\limits_{i=1}^{n} X_i$ 的估计中求 c, 使其在均方误差准则下最优;

(3) 判断由 (2) 得到的估计量是否为 θ 的相合估计量.

解: (1) 由题

$$\mu_1 = E(X) = \int_0^{+\infty} x \cdot \frac{1}{\theta}e^{-\frac{x}{\theta}}\mathrm{d}x \xlongequal{u=\frac{x}{\theta}} \int_0^{+\infty} \theta u e^{-u}\mathrm{d}u = \theta\Gamma(2) = \theta.$$

因此 θ 的矩估计量 $\hat{\theta}_1 = \mu_1 = E(\overline{X})$, 即样本均值为 θ 的矩估计量.

设 $x_1, x_2, \cdots, x_n$ 是样本 $X_1, X_2, \cdots, X_n$ 的一组观测值.

从而似然函数 $L(\theta) = \prod\limits_{i=1}^{n} f(x_i;\theta) = \begin{cases} \dfrac{1}{\theta^n}\prod\limits_{i=1}^{n}\mathrm{e}^{-\frac{x_i}{\theta}}, & x_i > 0, \\ 0, & \text{其他}. \end{cases}$

则对数似然函数 $l(\theta)=\ln L(\theta)=\begin{cases}-n\ln\theta-\dfrac{1}{\theta}\sum\limits_{i=1}^{n}x_i, & x_i>0,\\ 0, & 其他.\end{cases}$

从而 $x_i>0$ 时, $\dfrac{\mathrm{d}l(\theta)}{\mathrm{d}\theta}=-\dfrac{n}{\theta}+\dfrac{1}{\theta^2}\sum\limits_{i=1}^{n}x_i$.

令 $\left.\dfrac{\mathrm{d}l(\theta)}{\mathrm{d}\theta}\right|_{\theta=\hat{\theta}}=0$, 则 $-\dfrac{n}{\hat{\theta}}+\dfrac{1}{\hat{\theta}^2}\sum\limits_{i=1}^{n}x_i=0$, 解得 $\hat{\theta}=\dfrac{1}{n}\sum\limits_{i=1}^{n}x_i=\overline{x}$.

因此 θ 的极大似然估计量为 $\hat{\theta}=\overline{X}$, 也即样本均值为 θ 的极大似然估计量.

(2) 此时令 $\hat{\theta_0}=c\sum\limits_{i=1}^{n}X_i$.

由于

$$E(X^2)=\int_0^{+\infty}x^2\cdot\frac{1}{\theta}e^{-\frac{x}{\theta}}\mathrm{d}x\xlongequal{u=\frac{x}{\theta}}\int_0^{+\infty}\theta^2u^2e^{-u}\mathrm{d}u=\theta^2\Gamma(3)=2\theta^2.$$

则 $\mathrm{Var}(X)=E(X^2)-(E(X))^2=2\theta^2-\theta^2=\theta^2$.

因为 $E(\hat{\theta_0})=c\sum\limits_{i=1}^{n}E(X_i)=cnE(X)=nc\theta$; $\mathrm{Var}(\hat{\theta_0})=c^2\mathrm{Var}\left(\sum\limits_{i=1}^{n}X_i\right)=c^2n\mathrm{Var}(X)=nc^2\theta^2$.

因此 $E(\hat{\theta_0}^2)=(E(\hat{\theta_0}))^2+\mathrm{Var}(\hat{\theta_0})=c^2n^2\theta^2+nc^2\theta^2=n(n+1)c^2\theta^2$.

此时

$$\begin{aligned}\mathrm{Mse}(\hat{\theta_0})=E[(\hat{\theta_0}-\theta)^2]&=E(\hat{\theta_0}^2)-2\theta E(\hat{\theta_0})+\theta^2=n(n+1)c^2\theta^2-2nc\theta^2+\theta^2\\&=\theta^2[n(n+1)c^2-2nc+1].\end{aligned}$$

在均方误差准则下最优时, $\mathrm{Mse}(\hat{\theta_0})$ 最小, 即 $n(n+1)c^2-2nc+1$ 取最小值.

因此 $c=\dfrac{2n}{2n(n+1)}=\dfrac{1}{n+1}$ 时, 均方误差准则下最优.

(3) 由(2), $\hat{\theta_0}=\dfrac{1}{n+1}\sum\limits_{i=1}^{n}X_i$, 且有

$$E(|\hat{\theta_0}-\theta|^2)=E[(\hat{\theta_0}-\theta)^2]=\theta^2\left(\frac{n}{n+1}-\frac{2n}{n+1}+1\right)=\frac{\theta^2}{n+1}.$$

从而由马尔可夫不等式, 对任意 $\varepsilon>0$, 有

$$P\{|\hat{\theta_0}-\theta|<\varepsilon\}\geqslant 1-\frac{E(|\hat{\theta_0}-\theta|^2)}{\varepsilon^2}=1-\frac{\theta^2}{(n+1)\varepsilon^2}.$$

因为 $n\to+\infty$ 时, $1-\dfrac{\theta^2}{(n+1)\varepsilon^2}\to 1$, 且 $P\{|\widehat{\theta}_2-\theta|<\varepsilon\}\leqslant 1$.

由夹逼准则, $\lim\limits_{n\to+\infty}P\{|\hat{\theta_0}-\theta|<\varepsilon\}=1$, 从而 $\hat{\theta_0}\xrightarrow{P}\theta$.

即 $\hat{\theta_0}$ 是 θ 的相合估计量.

B7. 设总体 X 的密度函数为

$$f(x;\theta,\lambda)=\begin{cases}\dfrac{\lambda x^{\lambda-1}}{\theta^{\lambda}}, & 0<x<\theta,\\[2ex] 0, & 其他,\end{cases}$$

其中 $\theta>0,\lambda>0,X_1,X_2,\cdots,X_n$ 是来自总体 X 的简单随机样本.

(1) $\lambda=3,\theta$ 为未知参数, 求 θ 的矩估计量, 并判断其是否为 θ 的无偏估计, 说明理由;

(2) $\theta=3,\lambda$ 为未知参数, 求 λ 的极大似然估计量, 并判断其是否为 λ 的相合估计, 说明理由.

解: (1) 此时 $f(x;\theta)=\begin{cases}\dfrac{3x^2}{\theta^3}, & 0<x<\theta,\\[2ex] 0, & 其他.\end{cases}$ 则有

$$\mu_1=E(X)=\int_0^{\theta}x\cdot\frac{3x^2}{\theta^3}\mathrm{d}x=\frac{3}{4\theta^3}x^4\bigg|_0^{\theta}=\frac{3\theta}{4}.$$

此时 $\theta=\dfrac{4}{3}\mu_1$, 用 A_1 代替 μ_1, 则 θ 的矩估计量

$$\hat{\theta}=\frac{4}{3}A_1=\frac{4}{3}\overline{X}.$$

因为 $E(\hat{\theta})=\dfrac{4}{3}E(\overline{X})=\dfrac{4}{3}E(X)=\dfrac{4}{3}\cdot\dfrac{3\theta}{4}=\theta.$

从而矩估计量是 θ 的无偏估计.

(2) 此时 $f(x;\lambda)=\begin{cases}\dfrac{\lambda x^{\lambda-1}}{3^{\lambda}}, & 0<x<3,\\[2ex] 0, & 其他.\end{cases}$ 则似然函数 $L(\lambda)=\prod\limits_{i=1}^{n}f(x_i;\lambda)=\begin{cases}\dfrac{\lambda^n}{3^{n\lambda}}\prod\limits_{i=1}^{n}x_i^{\lambda-1}, & 0<x_i<3,\\[2ex] 0, & 其他.\end{cases}$

对数似然函数 $l(\lambda)=\begin{cases}n\ln\lambda-n\lambda\ln 3+(\lambda-1)\sum\limits_{i=1}^{n}\ln x_i, & 0<x_i<3,\\[2ex] 0, & 其他.\end{cases}$

$0<x_i<3$ 时 $\dfrac{\mathrm{d}l(\lambda)}{\mathrm{d}\lambda}=\dfrac{n}{\lambda}-n\ln 3+\sum\limits_{i=1}^{n}\ln x_i.$

令 $\dfrac{\mathrm{d}l(\lambda)}{\mathrm{d}\lambda}\bigg|_{\lambda=\hat{\lambda}}=0$, 则 $\dfrac{n}{\hat{\lambda}}-n\ln 3+\sum\limits_{i=1}^{n}\ln x_i=0$, 解得 $\hat{\lambda}=\dfrac{1}{\ln 3-\dfrac{1}{n}\sum\limits_{i=1}^{n}\ln x_i}.$

因此 λ 的极大似然估计量

$$\hat{\lambda}=\frac{1}{\ln 3-\dfrac{1}{n}\sum\limits_{i=1}^{n}\ln X_i}.$$

因为 $\lambda>0$ 时, $\lim\limits_{x\to 0}x^{\lambda}\ln x=0$, 则有

$$E(\ln X)=\int_0^3\ln x\cdot\frac{\lambda x^{\lambda-1}}{3^{\lambda}}\mathrm{d}x=\frac{1}{3^{\lambda}}\int_0^3\ln x\,\mathrm{d}x^{\lambda}=\frac{1}{3^{\lambda}}\left(x^{\lambda}\ln x\Big|_0^3-\int_0^3 x^{\lambda}\mathrm{d}\ln x\right)$$

$$
\begin{aligned}
&= \frac{1}{3^{\lambda}}\left(3^{\lambda}\ln 3 - 0 - \int_0^3 x^{\lambda-1}\mathrm{d}x\right) = \ln 3 - \frac{1}{\lambda\cdot 3^{\lambda}}x^{\lambda}\Big|_0^3 \\
&= \ln 3 - \frac{1}{\lambda\cdot 3^{\lambda}}(3^{\lambda}-0) = \ln 3 - \frac{1}{\lambda}.
\end{aligned}
$$

由辛钦大数定律

$$
\frac{1}{n}\sum_{i=1}^{n}\ln X_i \xrightarrow{P} E(\ln X) = \ln 3 - \frac{1}{\lambda}.
$$

因此 $\ln 3 - \frac{1}{n}\sum\limits_{i=1}^{n}\ln X_i \xrightarrow{P} \frac{1}{\lambda}$, 则有 $\hat{\lambda} \xrightarrow{P} \lambda$.

则 $\hat{\lambda}$ 为 λ 的相合估计.

B8. 设总体 X 的密度函数为

$$
f(x;\theta) = \begin{cases} \mathrm{e}^{-(x-\theta)}, & x \geqslant \theta, \\ 0, & \text{其他}, \end{cases}
$$

$\theta > 0$ 未知, $X_1, X_2, \cdots, X_n$ 是来自总体 X 的简单随机样本.

(1) 求 θ 的极大似然估计量 $\widehat{\theta}$;

(2) 求 $\widehat{\theta} - \theta$ 的密度函数;

(3) 判断 $\widehat{\theta} - \theta$ 是否可以取为关于 θ 的区间估计问题的枢轴量;

(4) 求 θ 的置信水平为 $1-\alpha$ 的单侧置信下限.

解: (1) 由题可得似然函数 $L(\theta) = \prod\limits_{i=1}^{n} f(x_i;\theta) = \begin{cases} \prod\limits_{i=1}^{n} e^{-(x_i-\theta)}, & x_i \geqslant \theta, \\ 0, & \text{其他}. \end{cases}$

则对数似然函数 $l(\theta) = \ln L(\theta) = \begin{cases} -\sum\limits_{i=1}^{n}(x_i-\theta) = n\theta - \sum\limits_{i=1}^{n} x_i, & x_i \geqslant \theta, \\ 0, & \text{其他}. \end{cases}$.

从而 $x_i \geqslant \theta$ 时 $\dfrac{\mathrm{d}l(\theta)}{\mathrm{d}\theta} = n > 0$.

即对数似然函数递增, 从而当 $l(\theta)$ 最大时, θ 要取到最大.

由题, $\theta \leqslant x_i$ 成立, 从而 $\theta \leqslant \min\{x_1, x_2, \cdots, x_n\}$. 则此时应取 $\theta = \min\{x_1, x_2, \cdots, x_n\}$.

因此 θ 的极大似然估计量

$$
\widehat{\theta} = \min\{X_1, X_2, \cdots, X_n\}.
$$

(2) 由于 X 的分布函数 $F(x)=\begin{cases}1-\mathrm{e}^{-(x-\theta)}, & x\geqslant\theta,\\ 0, & \text{其他}.\end{cases}$

对 $\widehat{\theta}=\min\{X_1,X_2,\cdots,X_n\}$, 其分布函数 $F_N(x)=1-(1-F(x))^n=\begin{cases}1-\mathrm{e}^{-n(x-\theta)}, & x\geqslant\theta,\\ 0, & \text{其他}.\end{cases}$

由于 $\widehat{\theta}-\theta$ 的分布函数

$$G(x)=P\{\widehat{\theta}-\theta\leqslant x\}=P\{\widehat{\theta}\leqslant x+\theta\}=F_N(x+\theta)=\begin{cases}1-e^{-nx}, & x\geqslant 0,\\ 0, & \text{其他}.\end{cases}$$

因此 $\widehat{\theta}-\theta$ 的密度函数

$$g(x)=G'(x)=\begin{cases}ne^{-nx}, & x\geqslant 0,\\ 0, & \text{其他}.\end{cases}$$

(3) 由(2)可知, $\widehat{\theta}-\theta$ 的分布已知, 且形式上不依赖其他未知参数.

从而 $\widehat{\theta}-\theta$ 可以作为区间估计问题的枢轴量.

(4) 取枢轴量 $\widehat{\theta}-\theta$, 因为 $t>0$ 时 $P\{\widehat{\theta}-\theta<t\}=1-e^{-nt}$.

从而令 $1-\alpha=P\{\widehat{\theta}-\theta<t\}$, 解得 $t=-\dfrac{1}{n}\ln\alpha$.

则有

$$P\{\widehat{\theta}-\theta<t\}=P\{\theta>\widehat{\theta}-t\}=1-\alpha.$$

则此时 θ 的置信水平为 $1-\alpha$ 的单侧置信下限为

$$\widehat{\theta}-t=\min\{X_1,X_2,\cdots,X_n\}+\frac{\ln\alpha}{n}.$$

B9. 假设某批日光灯的寿命服从正态分布 $N(\mu,1358)$, 从该总体中随机抽出 36 个个体, 计算得到置信水平为 $1-\alpha$ 的置信区间的长度为 5. 若保持置信水平不变, 要使得区间长度变为 2, 问样本容量应该为多少?

解:

B10. 某餐厅为了解顾客对餐厅新开发的菜品的满意程度, 随机调查来餐厅就餐的顾客 80 人, 结果发现有 55 人满意, 求满意比例 p 的置信水平为 95% 的置信区间.

解:

第 8 章　假设检验

A1. 一家大型超市接到许多消费者投诉某品牌袋装土豆片 (标注 60 克/袋) 的重量不符合标准, 为了维护消费者和供应商的利益, 超市管理员决定对下一批袋装土豆片的平均重量 μ (单位: g) 进行抽样检验, 提出如下原假设和备择假设:

$$H_0: \mu \geqslant 60, \quad H_1: \mu < 60.$$

(1) 分析这一假设检验问题的第 I 类错误和第 II 类错误;

(2) 从消费者和供应商的角度出发, 你认为他们分别希望控制犯哪类错误的概率?

解:

A2. 电视机显像管的质量标准是平均使用寿命为 15000 h. 某电视机厂宣称其生产的显像管平均寿命大大高于规定的标准. 为了对此说法进行验证, 随机抽取了 100 件该厂生产的显像管, 测得平均使用寿命为 15525 h. 假设该厂生产的显像管的寿命 $X \sim N(\mu, 1500^2)$, 利用假设检验推断是否有充分的理由认为该厂的显像管寿命显著地高于规定的标准 (显著性水平 $\alpha = 0.05$)

(1) 给出检验的原假设、备择假设、检验统计量和拒绝域, 并根据样本资料作出判断;

(2) 计算 $P-$ 值作出推断, 和 (1) 的判断结果是否一致?

解: (1) 由题, 原假设为 $H_0: \mu \leqslant 15000$, 备择假设为 $H_1: \mu > 15000$.

由于方差已知, 此时检验统计量 $Z = \dfrac{\overline{X} - 15000}{\frac{1500}{\sqrt{n}}}$. 因此对显著性水平 $\alpha = 0.05$, $z_{0.05} = 1.645$.

从而拒绝域

$$W = \{Z > z_{0.05}\} = \{Z > 1.645\}.$$

由题中样本可知, $n = 100$, $\bar{x} = 15525$, 代入检验统计量, 其值 $z_0 = \dfrac{15525 - 15000}{\frac{1500}{\sqrt{100}}} = \dfrac{525}{150} = 3.5$.

因为 $z_0 > 1.645$, 此时拒绝原假设, 即有充分的把握认为该厂的显像管寿命显著地高于规定的标准.

(2) 因为此时 $P-$ 值 $= 1 - \Phi(z_0) = 1 - 0.9998 = 0.0002 < 0.05$.

从而此时拒绝原假设, 与 (1) 的判断结果一致.

A3. 食品厂用自动装罐机装罐头食品, 每罐的标准重量为 500 g. 为了检测机器是否正常工作, 每隔一定的时间进行抽样检验现随机抽得 10 罐, 测得平均重量为 498 g, 标准差为 6.5 g. 假定罐头的重量服从正态分布, 利用假设检验推断机器的工作是否正常 ($\alpha = 0.02$).

(1) 给出检验的原假设、备择假设、检验统计量和拒绝域, 并根据样本资料作出判断;

(2) 计算 $P-$ 值并作出推断, 和 (1) 的判断结果是否一致?

解: (1) 由题, 此时假定罐头重量 X 服从分布 $N(500,\sigma^2)$, 此时 σ^2 未知.

从而原假设为 $H_0:\mu=500$, 备择假设为 $H_1:\mu\neq 500$.

由于方差未知, 此时检验统计量 $T=\dfrac{\overline{X}-500}{\frac{S}{\sqrt{n}}}$. 此时有 $T\sim t(n-1)$.

由样本知 $n=10$, 因此对显著性水平 $\alpha=0.02$, $t_{0.01}(9)=2.8214$.

从而拒绝域

$$W=\{|T|\geqslant t_{0.01}(9)\}=\{|T|\geqslant 2.8214\}.$$

由题中样本可知, $\overline{x}=498$, $s=6.5$. 代入检验统计量, 其值 $t_0=\dfrac{498-500}{\frac{6.5}{\sqrt{10}}}=\dfrac{-2\sqrt{10}}{6.5}=-0.973$.

因为 $|t_0|<2.8214$, 此时不能拒绝原假设, 即认为机器的工作正常.

(2) 利用 EXCEL 中 T.DIST 函数, 此时 $P-$ 值 $=2P\{t(9)\geqslant|t_0|\}=2\times 0.178=0.356>0.02$.

此时不能拒绝原假设, 与 (1) 的判断结果一致.

A4. 某高校管理部门为了解在校学生外卖消费情况, 随机抽查 100 名学生, 询问他们上个月的外卖消费额, 计算得平均消费额为 478 元, 消费标准差为 85 元. 设该校学生外卖消费额服从正态分布 $N(\mu,\sigma^2)$, 在显著性水平 $\alpha=0.05$ 下, 检验如下假设:

$$H_0:\mu\leqslant 450, H_1:\mu>450.$$

请给出检验统计量, 计算 $P-$ 值并作出推断.

解:

A5. 一调查咨询公司对某论坛的日发帖量进行为期 2 周的调查, 记录每天的日发帖量, 计算得 14 天的日平均发帖量为 506 条, 标准差为 6.26 条. 假定该论坛的日发帖量服从正态分布, 问是否有充分的理由认为该论坛的日平均发帖量小于 510 条 $(\alpha=0.05)$?

解:

A6. 某汽车厂商宣称他们生产的汽车平均每公升汽油可行驶 15 km 以上. 为验证该广告的真实性, 随机选取 10 辆汽车进行测试, 记录每辆车每公升汽油行驶的里程数, 得到如下数据:

14.8 15.1 16.9 14.8 13.7 12.9 13.5 14.9 15.4 13.5

假设数据服从正态分布, 利用假设检验分析该广告的可靠性 ($\alpha = 0.05$).

解:

A7. 根据《中国居民营养与慢性病状况报告 (2015) 》, 全国 18 岁及以上成年男性的平均身高为 1.67 m, 现从我国某地区随机抽选 400 名成年男子, 测得身高的平均值为 1.69 m, 标准差为 0.042 m. 设该地区男子身高服从 $N(\mu, \sigma^2)$, 问该地区男子的身高是否显著高于全国平均水平 ($\alpha = 0.05$)?

解:

A8. 为了解某犬类疫苗注射后是否会使得犬的体温升高, 随机选择 9 只狗, 记录它们注射疫苗前后的体温 (单位: °C):

编号	1	2	3	4	5	6	7	8	9
注射前体温/°C	37.5	37.7	38.1	37.9	38.3	38.5	38.1	37.6	38.4
注射后体温/°C	37.7	38.0	38.2	37.9	38.2	38.8	38.0	37.5	38.8

设注射疫苗前后体温差服从正态分布, 试用成对数据的假设检验推断是否有充足的理由认为狗注射该疫苗后体温显著升高 ($\alpha = 0.05$).

解: 设注射前后体温分别为 X_i, Y_i, 差值 $D_i = Y_i - X_i$, $i = 1, 2, \cdots, 9$.

此时 D_i 可看成来自正态总体 $N(\mu_D, \sigma_D^2)$ 的样本, σ_D^2 未知.

考虑右侧检验

$$H_0 : \mu_D \leqslant 0, \quad H_1 : \mu_D > 0.$$

对 $\overline{D} = \dfrac{1}{n}\sum_{i=1}^{n} D_i$, $S_D^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}(D_i - \overline{D})^2$, 此时检验统计量

$$T = \frac{\overline{D}}{\frac{S_D}{\sqrt{n}}} = \frac{\sqrt{n}\overline{D}}{S_D}.$$

由样本可知, $n = 9$, $\overline{d} = \dfrac{1}{9}(0.2 + 0.3 + 0.1 + 0 - 0.1 + 0.3 - 0.1 - 0.1 + 0.4) = \dfrac{1}{9} = 0.1111$,

$s_d^2 = \dfrac{1}{8}\left(0.2^2 + 0.3^2 + 0.1^2 + 0^2 + 0.1^2 + 0.3^2 + 0.1^2 + 0.1^2 + 0.4^2 - 9 \times \dfrac{1}{81}\right) = 0.0386.$

从而 $s_d = 0.1965$, 代入检验统计量有 $t_0 = \dfrac{3 \times 0.1111}{0.1965} = 1.696$.

查表有 $t_{0.05}(8) = 1.8595$, 因此检验的拒绝域

$$W = \{T \geqslant t_{0.05}(8)\} = \{T \geqslant 1.8595\}.$$

因为 $t_0 < 1.8595$, 从而不能拒绝原假设, 选择接受 H_0.

即认为狗注射该疫苗后体温没有显著升高.

A9. 一减肥药广告宣称: 减肥者服用 2 周后, 体重会明显下降. 消费者协会为了对该减肥药的减肥效果进行评估, 随机抽选 10 位服用该减肥药的顾客, 记录其服用减肥药前和服用减肥药 2 周后的体重 (单位: kg):

编号	1	2	3	4	5	6	7	8	9	10
服药前体重/kg	66	70	56	58	49	75	63	56	48	75
服药后体重/kg	68	65	54	59	45	70	60	50	47	68

设服药前后体重差服从正态分布, 试用成对数据的假设检验分析该广告是否可靠 ($\alpha = 0.05$).

解:

A10. 某经销代理商和乳业公司的合约里要求 225 mL 盒装牛奶的容量标准差不可超过 8 mL, 否则就予以退货. 现随机抽取 15 盒牛奶, 测得容量 (单位: mL) 如下:

230 223 228 229 220 215 217 231

220 223 230 224 226 228 227

假设样本来自正态总体 $N(\mu, \sigma^2)$, 其中 $-\infty < \mu < +\infty, \sigma > 0$ 均未知. 在显著性水平 $\alpha = 0.05$ 下, 通过计算 $P-$ 值来检验假设

$$H_0 : \sigma \geqslant 8, H_1 : \sigma < 8.$$

解:

A11. 已知某厂生产的某种零件的长度 (单位: cm) 服从正态分布, 要求零件的标准长度为 15 cm, 标准差不超过 0.2 cm. 现从该厂中随机抽取 16 只零件, 测得长度如下:

15.1 14.9 14.8 14.6 15.2 14.8 14.9 14.6

14.8 15.1 15.3 14.7 15.0 15.2 15.1 14.7

利用假设检验推断这批零件是否符合标准要求 ($\alpha = 0.05$).

解:

A12. 下列数据为 A, B 两个煤矿开采的每吨煤产生的热量记录 (单位: 4.186×10^3 J):

A 矿	8 500	8 330	8 480	7 960	8 030
B 矿	7 710	7 890	7 920	8 270	7 860

假设样本来自两个方差相等且相互独立的正态总体, 是否可以认为 A 矿的煤产生的热量要显著地大于 B 矿的煤 ($\alpha = 0.05$)?

解:

A13. 为比较甲、乙两位电脑打字员的出错情况, 随机抽查甲输入的文件 8 页, 各页出错字数为

5 3 2 0 1 2 2 4

随机抽查乙输入的文件 9 页, 各页出错字数为

5 1 3 2 4 6 4 2 5

假设甲、乙两人页出错字数都服从正态分布.

(1) 检验甲、乙两人页出错数的方差是否相等 ($\alpha = 0.05$);

(2) 根据(1)的检验结果选择合适的检验方法, 推断打字员甲的平均页出错字数是否显著少于打字员乙 ($\alpha = 0.05$).

解:

A14. 为了研究男性长跑运动员的心率是否低于一般健康男性, 现从省长跑队随机抽取 10 名男运动员, 从某高校随机抽取 25名健康状况良好的男学生. 测得运动员心率的平均值为 60 次/分, 标准差为 6 次/分, 大学生心率的平均值为 73 次/分, 标准差为 13 次/分. 假设心率服从正态分布.

(1) 根据上面的资料检验两个群体心率的方差是否相等 ($\alpha = 0.05$);

(2) 根据 (1) 的检验结果选择合适的检验方法推断是否有充分的理由认为男性长跑运动员的心率显著低于一般健康男性 ($\alpha = 0.05$).

解:

B1. 设总体 $X \sim N(\mu, \sigma^2)$, 从总体中抽取样本容量为 16 的简单随机样本, 样本均值为 $\overline{X}$, 样本方差为 S^2.

(1) 若 $\sigma^2 = 1$, 在显著性水平为 0.05 下对于假设: $H_0 : \mu = 1$, $H_1 : \mu \neq 1$, 给出检验统计量和拒绝域, 并计算在 $\mu = 2$ 时犯第 II 类错误的概率;

(2) 若 μ 未知, 在显著性水平为 0.05 下对于假设: $H_0 : \sigma^2 = 1$, $H_1 : \sigma^2 > 1$, 给出检验统计量和拒绝域, 并计算在 $\sigma^2 = 4$ 时犯第 II 类错误的概率;

(3) 若根据样本值计算得 $\overline{x} = 1.54$, $s^2 = 1.44$, 求 (1) 和 (2) 中的 $P-$ 值.

解: (1) 由题, $\sigma^2 = 1$ 已知. 且样本容量 $n = 16$. 此时取检验统计量 $Z = \dfrac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} = 4(\overline{X} - 1)$.

查表有 $z_{0.025} = 1.96$, 从而在显著性水平 0.05 下, 拒绝域

$$W = \{|Z| \geqslant z_{0.025}\} = \{4|\overline{X} - 1| \geqslant 1.96\}.$$

$\mu = 2$ 时犯第 II 类错误, 即此时接受 H_0, 有 $4|\overline{X} - 1| < 1.96$, 则 $0.51 < \overline{X} < 1.49$.

因为 $\dfrac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0, 1)$, 从而所求概率

$$\beta = P\{\text{接受}H_0|\mu = 2\} = P\{0.51 < \overline{X} < 1.49|\mu = 2\} = P\left\{\frac{0.51 - 2}{\frac{1}{\sqrt{16}}} < \frac{\overline{X} - 2}{\frac{1}{\sqrt{16}}} < \frac{1.49 - 2}{\frac{1}{\sqrt{16}}}\right\}$$

$$= \Phi(-2.04) - \Phi(-5.96) = 1 - \Phi(2.04) - (1 - \Phi(5.96)) = \Phi(5.96) - \Phi(2.04)$$

$$= 1 - 0.9793 = 0.0207.$$

(2) 因为 $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$, 由题对 $n = 16$, $\sigma^2 = 1$ 时取检验统计量 $\chi^2 = 15S^2$.

因为 $H_1 : \sigma^2 > 1$, 查表有 $\chi^2_{0.05}(15) = 24.996$, 从而在显著性水平 0.05 下, 拒绝域

$$W = \{\chi^2 \geqslant \chi^2_{0.05}(15)\} = \{15S^2 \geqslant 24.996\}.$$

$\sigma^2 = 4$ 时犯第 II 类错误, 即此时接受 H_0, 有 $15S^2 < 24.996$.

因为 $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$, 从而所求概率

$$\beta = P\{\text{接受}H_0|\sigma^2 = 4\} = P\{15S^2 < 24.996|\sigma^2 = 4\} = P\left\{\frac{15S^2}{4} < 6.249\right\}$$

$$= 1 - P\{\chi^2(15) \geqslant 6.249\} = 1 - 0.9753 = 0.0247.$$

其中 $P\{\chi^2(15)\geqslant 6.249\}$ 是借助 EXCEL 中 CHISQ.DIST.RT 函数计算得到, 下同.

(3) 对(1), 将样本值代入, 则检验统计量的值为 $z_0=4(\overline{x}-1)=2.16$. 此时

$$P\_\text{值}=P_{H_0}\{|Z|\geqslant|z_0|\}=2(1-\Phi(|z_0|)=2-2\Phi(2.16)=2-2\times0.9846=0.0308.$$

对(2), 将样本值代入, 则检验统计量的值为 $\chi_0^2=15s^2=21.6$. 此时

$$P-\text{值}=1-P\{\chi^2(15)\leqslant\chi_0^2\}=P\{\chi^2(15)>21.6\}=0.1187.$$

B2. 对选择去英语国家继续深造的高校毕业生而言, 为了能申请到心仪的学校和更好地适应新环境的学习及生活, 需要尽可能地提高自己的英语水平. 某英语培训机构宣称, 参加该机构开办的为期 4 周的一对一培训课程后, TOEFL (托福) 平均成绩可提升 7 分. 一市场调查公司为了对该机构的培训效果进行评估, 随机选择了 12 位参加过该机构一对一培训课程的学生, 了解他们培训前后的 TOEFL 成绩, 具体如下:

编号	1	2	3	4	5	6	7	8	9	10	11	12
培训前 TOEFL 成绩	76	85	78	90	104	87	91	83	95	108	93	84
培训后 TOEFL 成绩	89	92	90	93	106	96	100	90	100	110	100	95

设培训前后成绩差服从正态分布, 试用成对数据的假设检验分析该培训机构的宣称是否可靠 ($\alpha=0.05$).

解:

B3. 为了比较 A 高校和 B 高校教师的收入水平, 在两所高校具有副高职称的教师中各随机选择了 36 位和 49 位年龄在 40 ~ 50 岁的教师, 了解他们上一年的税前工资收入. 计算得 A 高校教师的平均年工资收入为 28.8 万元, 标准差为 11.6 万元; B 高校教师的平均年工资收入为 22.4 万元, 标准差为 8.4 万元.

请选择合适的假设检验方法对下列问题进行推断:

(1) 从收入的差距程度来看, B 高校教师的收入差距程度是否显著低于 A 高校 ($\alpha=0.05$)?

(2) 从收入的平均水平来看, 是否有充分的理由认为 A 高校教师的平均收入要比 B 高校至少多 5 万元 ($\alpha=0.1$)?

解:

B4. 火药生产厂家设计出一种新的火药生产方案, 要求使子弹发射的枪口速度为 900 m/s. 假设枪口速度

X (单位: m/s) 服从正态分布 $N(\mu,\sigma^2)$, 现做了 8 次试验, 其速度分别为

893 886 897 903 901 898 909 889

(1) 求平均速度 μ 的置信水平为 95% 的置信区间;

(2) 根据区间估计的结果, 推断子弹发射的枪口速度是否与设计要求有显著差异;

(3) 利用假设检验, 在显著性水平 $\alpha=0.05$ 下, 计算 $P-$ 值并推断子弹发射的枪口速度是否与设计要求有显著差异; 结果和(2)是否一致?

解:

B5. 某个八面体各面分别标有数字 1,2,3,4,5,6,7,8. 为检验各面是否匀称, 即各面出现的概率是否相等; 作 600 次投掷试验, 各数字朝上的次数如下:

数字	1	2	3	4	5	6	7	8
频数	72	83	78	90	70	71	64	72

在显著性水平 0.05 下, 检验假设 H_0: 该八面体是匀称的

解:

B6. 对某公交车站观察从 12 点到 15 点这 3 个小时前来等车的乘客情况, 将 2 min 作为一个单位时间, 记录 90 个单位时间内的等车人数, 数据如下:

人数	0	1	2	3	4	5	6	7	>7
频数	5	12	18	21	16	13	3	2	0

试利用上述数据推断在该公交车站的候车人数是否服从泊松分布 ($\alpha=0.05$).

解:

B7. 一盒中有 10 个球, 其中红球有 a 个 (未知), 其余是白球, 采用有放回抽样取 3 个球作为一次实验, 这样的试验总共进行 200 次, 发现有 40 次没有取到红球, 有 85 次取到 1 个红球, 有 63 次取到2个红球, 有 12 次取到 3 个红球, 在显著性水平 $\alpha=0.05$ 下检验假设 $H_0: a=3$.

解:

B8. 设 X 为前后两位客户到某自动取款机办理业务的时间间隔 (单位: min), 现观察了 120 次, 获得如下

数据:

等候时间	$0 \leqslant x \le 5$	$5 < x \leqslant 10$	$10 < x \leqslant 20$	$20 < x \leqslant 30$	$x > 30$
频数	45	27	25	12	11

在显著性水平 $\alpha = 0.05$ 下, 检验假设 H_0: 前后两位客户的时间间隔服从均值为10的指数分布.

解:

B9. 对某地区成年男子的身高 X (单位: cm) 进行观察, 随机抽取 200 名男子, 得到样本均值和样本标准差分别为 $\overline{x} = 169.9$, $s = 9.6$, 其他资料如下:

身高	$x \leqslant 163$	$163 < x \leqslant 167$	$167 < x \leqslant 171$	$171 < x \leqslant 175$	$175 < x \leqslant 179$	$179 < x \leqslant 183$	$x > 183$
频数	41	34	40	33	27	16	9

在显著性水平 $\alpha = 0.05$ 下, 检验假设 H_0: 该地区成年男子的身高服从正态分布.

解:

B10. 考察某特定人群, 其收入与文化消费支出有无关联. 把收入分为低、中、高三档, 文化消费支出分成低、高两档. 从中随机抽取 200 人, 得结果如下:

收入	文化消费支出		合计
	低	高	
低	64	16	80
中	36	16	52
高	60	8	68
合计	160	40	200

在显著性水平 $\alpha = 0.05$ 下, 检验假设

$$H_0\text{: 收入与文化消费支出相互独立.}$$

解: