Probability and Statistics Note

概率论与数理统计笔记

Qui non est hodie cras minus aptus erit

整理：jjyu

整理时间：December 25, 2018

Email: yus@sustc.edu.cn

Version: 1.00

目　录

第 1 章
第一章：概率

Remark: 如同高数或者高中应用题一样，题目中如果没有给定，一般来说，解决问题的第一步是定义事件或者随机变量是什么，然后再根据题意进行运算!!!!

1.1 定义和概型计算

1. 所有可能的实验结果的全体称为该实验的样本空间，记为 Ω 其中的元素记为 ω. 满足一定条件的样本点的集合称为事件。

2. 事件的基本运算：交并补，如何描述。事件运算的性质：交换律，结合律，分配律，德摩根律。

3. 概率测度是定义在样本空间 Ω 子集上的实函数，且满足

   - 规范（归一）性：$\mathbb{P}(\Omega) = 1$;
   - 非负性：$\mathbb{P}(A) \geq 0$;
   - 可列可加性（σ-可加性）：对两两不相容的事件列 $\{A_k\}_{k=1}^{\infty}$, 有

$$\mathbb{P}(\cup_{k=1}^{\infty} A_k) = \sum_{k=1}^{\infty} \mathbb{P}(A_k).$$

4. 概率测度的基本性质

   - $\mathbb{P}(\emptyset) = 0$;
   - 单调性：如果 $A \subset B$, 那么 $\mathbb{P}(A) \leq \mathbb{P}(B)$;
   - 有限可加性：对两两不相容的事件列 $A_1, \cdots, A_n$, 有

$$\mathbb{P}(\cup_{k=1}^{n} A_k) = \sum_{k=1}^{n} \mathbb{P}(A_k).$$

5. 加法公式——计算事件和的概率：$\mathbb{P}(A\cup B)=\mathbb{P}(A)+\mathbb{P}(B)-\mathbb{P}(A\cap B)$, 更一般的，对任意事件列 $A_1,\cdots,A_n$, 有

$$\begin{aligned}\mathbb{P}(\cup_{k=1}^n A_k) &= \sum_{k=1}^n \mathbb{P}(A_k)-\sum_{1\le i<j\le n}\mathbb{P}(A_iA_j)\\ &+\sum_{1\le i<j<k\le n}\mathbb{P}(A_iA_jA_k)\\ &+\cdots\\ &+(-1)^{n-1}\mathbb{P}(A_1\cdots A_n).\end{aligned}$$

6. Remark: 加法公式对有限事件列成立，但对可列不一定成立。

7. 概型计算：古典概型——样本点是等可能的；加法原理和乘法原理。

例 1.1 从 $1,\cdots,2000$ 中随机取一个数，问取到的整数既不能被 6 整除也不能被 8 整除的概率是多少?

解答：

定义事件 $A:=\{$随机取出的数能被 6 整除$\}$ 和 $B:=\{$随机取出的数能被 6 整除$\}$, 则事件 $C:=\{$随机取出的数既不能被 6 整除也不能被 8 整除$\}$. 则可知 $C=\bar{A}\bar{B}$, 由德摩根律可知 $C=\Omega-A\cup B$. 因此

$$\begin{aligned}\mathbb{P}(C) &= 1-\mathbb{P}(A\cup B)\\ &= 1-\mathbb{P}(A)-\mathbb{P}(B)+\mathbb{P}(AB)\ (\text{加法公式})\\ &= 1-[2000/6]/2000-125/2000+[2000/24]/2000\end{aligned}$$

这里 $[\cdot]$ 是取整函数。

例 1.2 从 n 双尺码不同的鞋子中任取 $2r$, $(2r<n)$ 只，求下列事件的概率：

1. 所取的 $2r$ 只鞋子没有两只成对；
2. 所取的 $2r$ 只鞋子只有两只成对；
3. 所取的 $2r$ 只鞋子恰好做成 r 对.

解答：

从 $2n$ 只鞋子中任取 $2r$ 只总共有 $\binom{2n}{2r}$ 种取法：1、所取的 $2r$ 只鞋子没有两只成对意味着所取的鞋子分别属于 $2r$ 双鞋，因此可以先抽取 $2r$ 双鞋，然后每双鞋选取一只，总共有 $2^{2r}\binom{n}{2r}$, 所以事件的概率为 $\frac{2^{2r}\binom{n}{2r}}{\binom{2n}{2r}}$; 2、所取的 $2r$ 只鞋子只有两只成对，相当于从 n 双鞋中选取一双，然后剩下的选取的 $2r-2$ 只鞋子没有两只成对，总共有 $2^{2r-2}n\binom{n-1}{2r-2}$, 因此概率为 $\frac{2^{2r-2}n\binom{n-1}{2r-2}}{\binom{2n}{2r}}$. 3、所取的 $2r$ 只鞋子恰好做成 r 对相当于从 n 双鞋中选取 r 双，因此事件的概率为 $\frac{\binom{n}{r}}{\binom{2n}{2r}}$.

例 1.3 将 4 把能打开 4 间不同房门的钥匙随机发给 4 个人，试求至少有一个人能打开房门的概率。

解答：

记事件 $A_i := \{\text{第 i 个人能打开房门}\}, i = 1, 2, 3, 4$, 事件 $B := \{\text{至少有一个人能打开房门}\}$. 则有 $B = \cup_{i=1}^4 A_i$, 因此

$$\begin{aligned}
\mathbb{P}(B) &= \mathbb{P}(\cup_{i=1}^4 A_i) \\
&= \sum_{i=1}^4 \mathbb{P}(A_i) - \sum_{1\le i<j\le 4} \mathbb{P}(A_iA_j) \\
&\quad + \sum_{i<j<k} \mathbb{P}(A_iA_jA_k) - \mathbb{P}(A_1A_2A_3A_4) \\
&= 1 - 1/2! + 1/3! - 1/4!
\end{aligned}$$

第二个等号是利用加法公式。事件 A_i 意味着第 i 人拿对钥匙，其他三把钥匙任意分配，因此 $\mathbb{P}(A_i) = 3!/4!$, 同理，$\mathbb{P}(A_iA_j) = 2!/4!$ 和 $\mathbb{P}(A_iA_jA_k) = \mathbb{P}(A_1A_2A_3A_4) = 1$.

例 1.4 投掷六颗均匀骰子，结果中恰好出现四个不同数字的概率是多少？

解答：

投掷六颗骰子总的结果数是 6^6 种解法一、六颗骰子出现四个不同数字意味着有一个数字出现了三次或者两个数字：对于第一种情况：六颗骰子有三颗出现同一个数字，然后出现四个不同数字，因此对应的结果数为 $\binom{6}{3}\frac{6!}{2!}$. 对于第二种情况六颗骰子中有两颗出现相同数字，然后剩下四颗仍有两颗出现不同数字，总共有 $\frac{\binom{6}{2}\binom{4}{2}}{2}\frac{6!}{2!}$, 因此出现四个不同数字的概率为 $\frac{\binom{6}{3}\frac{6!}{2!}+\frac{\binom{6}{2}\binom{4}{2}}{2}\frac{6!}{2!}}{6^6}$.

解法二、第一种情况，首先从六个数字中选取四个数字，四个数字中选取一个数字重复三次，然后把六个数字分配给六颗骰子，总共有 $\binom{6}{4}\binom{4}{1}\frac{6!}{3!}$. 、第二种情况，首先从六个数字中选取四个数字，四个数字中选取 2 个数字重复两次，然后把六个数字分配给六颗骰子，总共有 $\binom{6}{4}\binom{4}{2}\frac{6!}{2!2!}$. 因此出现四个不同数字的概率为 $\frac{\binom{6}{4}\binom{4}{1}\frac{6!}{3!}+\binom{6}{4}\binom{4}{2}\frac{6!}{2!2!}}{6^6}$.

例 1.5 试证明：$P(A_1 \cup \cdots \cup A_n) \ge P(A_1) + \cdots + P(A_n) - (n-1)$. （证明题的基本步骤）

解答：

证明：（数学归纳法）当 $n = 2$ 时，$\mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2) - \mathbb{P}(A_1A_2) \ge \mathbb{P}(A_1) + \mathbb{P}(A_2) - 1$. 假设当 $n = k$, 结论成立。当 $n = k+1$ 时，记事件 $B = A_1 \cup \cdots \cup A_k$, 则

$$\begin{aligned}
\mathbb{P}(A_1 \cup \cdots \cup A_{k+1}) &= \mathbb{P}(B \cup A_{k+1}) \\
&\ge \mathbb{P}(A_1 \cup \cdots \cup A_k) + \mathbb{P}(A_{k+1}) - 1 \\
&\ge \sum_{i=1}^{k+1} \mathbb{P}(A_i) - k
\end{aligned}$$

因此当 $n = k+1$ 时，结论成立。由归纳法原理，结论成立。

1.2 独立性和条件概率

1. 给定事件 A, B, 且满足 $\mathbb{P}(B) > 0$. 则记

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(AB)}{\mathbb{P}(B)}$$

为在事件 B 发生的条件下，事件 A 发生的条件概率。

2. 性质：当 $\mathbb{P}(B) > 0$ 时这个定义才有意义；$\mathbb{P}(A|B) \geq \mathbb{P}(AB)$; 条件概率满足概率测度的三个条件：规范性，非负性，可列可加性。

3. 乘法公式: 给定事件 A, B, 且满足 $\mathbb{P}(B) > 0$. 则由定义可得

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

推广到 n 个事件的情形：

$$\begin{aligned}
\mathbb{P}(A_1A_2\cdots A_n) &= \mathbb{P}(A_n|A_1A_2\cdots A_{n-1})\mathbb{P}(A_1A_2\cdots A_{n-1}) \\
&= \cdots \\
&= \mathbb{P}(A_n|A_1A_2\cdots A_{n-1})\cdots\mathbb{P}(A_2|A_1)\mathbb{P}(A_1).
\end{aligned}$$

4. 全概率公式：事件 $B_1, B_2, \cdots, B_n$ 是样本空间的一个分划，则事件 $A = (AB_1) \cup (AB_2) \cup \cdots \cup (AB_n)$. 利用乘法公式可得

$$\mathbb{P}(A) = \mathbb{P}(A|B_1)\mathbb{P}(B_1) + \cdots + \mathbb{P}(A|B_n)\mathbb{P}(B_n).$$

5. Bayes 公式: 事件 $B_1, B_2, \cdots, B_n$ 是样本空间的一个分划，结合乘法公式和全概率公式可得

$$\mathbb{P}(B_i|A) = \frac{\mathbb{P}(A|B_i)\mathbb{P}(B_i)}{\sum\limits_{k=1}^{n}\mathbb{P}(A|B_k)\mathbb{P}(B_k)}.$$

6. **Remark**：什么时候用定义什么时候用 Bayes 公式，如果概率 $\mathbb{P}(B)$ 很容易求，则可以直接用定义，反之，可以用 Bayes 公式。

例 1.6 【三扇门问题】问题见课件。

解答： 假设事件 $A_i := \{\text{第 i 扇门后面是车}\}, i = 1, 2, 3$, 事件

$$B := \{\text{选择第一扇门后，主持人打开第二扇门}\}.$$

则由题意可知 $\mathbb{P}(A_1) = \mathbb{P}(A_2) = \mathbb{P}(A_3) = 1/3$, 利用全概率公式可知

$$\begin{aligned}
\mathbb{P}(B) &= \mathbb{P}(B|A_1)\mathbb{P}(A_1) + \mathbb{P}(B|A_2)\mathbb{P}(A_2) + \mathbb{P}(B|A_3)\mathbb{P}(A_3) \\
&= 1/2 * 1/3 + 0 * 1/3 + 1 * 1/3 = 1/2
\end{aligned}$$

因此，由 Bayes 公式

$$\begin{aligned}\mathbb{P}(A_1|B) &= \frac{\mathbb{P}(B|A_1)\mathbb{P}(A_1)}{\mathbb{P}(B|A_1)\mathbb{P}(A_1)+\mathbb{P}(B|A_2)\mathbb{P}(A_2)+\mathbb{P}(B|A_3)\mathbb{P}(A_3)} \\ &= \frac{1/2*1/3}{1/2} = 1/3.\end{aligned}$$

同理，$\mathbb{P}(A_3|B) = 2/3$. 因此，当主持人打开第二扇门后，更换选项更有利。

例 1.7 该习题介绍一个简单气象模型，更复杂的模型参见气象学文献。考虑连续几天的天气，令 R_i 表示第 i 天下雨这个事件。假设 $\mathbb{P}(R_i|R_{i-1}) = \alpha$ 和 $\mathbb{P}(R_i^c|R_{i-1}^c) = \beta$. 进一步假设只有今天的天气和明天的天气预报有关，即 $\mathbb{P}(R_i|R_{i-1}\cdots R_0) = \mathbb{P}(R_i|R_{i-1})$.

1. 如果今天下雨的概率是 p, 那么明天下雨的概率是多少?
2. 后天下雨的概率是多少?
3. n 天之后下雨的概率是多少? n 趋于无穷又会怎样?

解答：

1、由全概率公式可知

$$\begin{aligned}\mathbb{P}(R_{i+1}) &= \mathbb{P}(R_{i+1}|R_i)\mathbb{P}(R_i) + \mathbb{P}(R_{i+1}|R_i^c)\mathbb{P}(R_i^c) \\ &= \alpha p + (1-\beta)(1-p)\end{aligned}$$

2、由全概率公式可知

$$\begin{aligned}\mathbb{P}(R_{i+2}) &= \mathbb{P}(R_{i+2}|R_{i+1})\mathbb{P}(R_{i+1}) + \mathbb{P}(R_{i+2}|R_{i+1}^c)\mathbb{P}(R_{i+1}^c) \\ &= \alpha(\alpha p + (1-\beta)(1-p)) + (1-\beta)(1-\alpha p + (1-\beta)(1-p))\end{aligned}$$

3、由全概率公式，n 天之后下雨的概率 p_n 满足 $p_n = \alpha p_{n-1} + (1-\beta)(1-p_{n-1}) = (\alpha+\beta-1)p_{n-1} + (1-\beta)$. 如果 $\alpha+\beta = 1$, 则 $p_n = p_{n-1} = \cdots = p$;
如果 $a = \alpha+\beta-1 \neq 0$, 令 $b = 1-\beta$, 则 $p_n = ap_{n-1}+b$ 则利用递推可知

$$\mathbb{P}(R_{i+n}) = \frac{b}{1-a} + (p - \frac{b}{1-a})a^{n-1}.$$

则

$$\lim_{n\to\infty}\mathbb{P}(R_{i+n}) = \frac{1-\beta}{\alpha+\beta}.$$

例 1.8 很多人类疾病是遗传的（例如，血友病和泰萨二氏病），这是此类疾病的一个简单模型。基因型 aa 是有病的，在交配之前死亡。基因型 Aa 是一个携带者，基因型 AA 不是携带者，也没有病。

1. 如果两个携带者交配，他们的后代是这三种基因型之一的概率是多少?
2. 如果两个携带者的男性后代没有疾病，他是疾病携带者的概率是多少?

3. 假设第二问中的无病后代和没有家族病史的个体交配，并设其配偶是病毒携带者的概率是 p, 那么他们的后代是 AA, Aa, aa 的概率是多少?

4. 如果上一问的第一代没有疾病，那么基于此证据，其父辈是病毒携带者的概率是多少?

解答：

1、Aa 和 Aa 进行交配后代可能的基因型为 AA, Aa, aa 概率分别为 $P(AA) = P(aa) = 1/4, P(Aa) = 1/2$.

2、两个携带者的男性后代没有疾病的概率为 $P(AA) + P(Aa) = 3/4$, 则在后代没有疾病的条件下是携带者的概率为 $\frac{P(Aa)}{P(AA)+P(Aa)} = 2/3$.

3、假设事件 $A_1 = \{$无病后代基因为 AA$\}, A_2 = \{$无病后代基因为 Aa$\}$ 和事件 $B_1 = \{$配偶基因为 AA$\}, B_2 = \{$配偶基因为 Aa$\}$. 则由全概率公式

$$\begin{aligned}
\mathbb{P}(AA) &= \mathbb{P}(AA|A_1B_1)\mathbb{P}(A_1B_1) + \mathbb{P}(AA|A_1B_2)\mathbb{P}(A_1B_2) \\
&\quad + \mathbb{P}(AA|A_2B_1)\mathbb{P}(A_2B_1) + \mathbb{P}(AA|A_2B_2)\mathbb{P}(A_2B_2) \\
&= 1/3(1-p) + 1/6p + 1/3(1-p) + 1/6p = 2/3 - 1/3p.
\end{aligned}$$

利用相同的方法可得

$$\begin{aligned}
\mathbb{P}(Aa) &= \mathbb{P}(Aa|A_1B_1)\mathbb{P}(A_1B_1) + \mathbb{P}(Aa|A_1B_2)\mathbb{P}(A_1B_2) \\
&\quad + \mathbb{P}(Aa|A_2B_1)\mathbb{P}(A_2B_1) + \mathbb{P}(Aa|A_2B_2)\mathbb{P}(A_2B_2) \\
&= 0 + 1/6p + 1/3(1-p) + 1/3p = 1/3 + 1/6p.
\end{aligned}$$

和

$$\begin{aligned}
\mathbb{P}(aa) &= \mathbb{P}(aa|A_1B_1)\mathbb{P}(A_1B_1) + \mathbb{P}(aa|A_1B_2)\mathbb{P}(A_1B_2) \\
&\quad + \mathbb{P}(aa|A_2B_1)\mathbb{P}(A_2B_1) + \mathbb{P}(aa|A_2B_2)\mathbb{P}(A_2B_2) \\
&= 0 + 0 + 0 + 1/6p = 1/6p.
\end{aligned}$$

4、则假设事件 $B = \{$父辈是携带者$\}$ 和 $A = \{$第一代没有疾病$\}$, 则

$$\begin{aligned}
\mathbb{P}(B|A) &= \frac{\mathbb{P}(AB)}{\mathbb{P}(A)} \\
&= \frac{\mathbb{P}(AA|A_2B_1)\mathbb{P}(A_2B_1) + \mathbb{P}(AA|A_2B_2)\mathbb{P}(A_2B_2)}{\mathbb{P}(AA) + \mathbb{P}(Aa)} \\
&\quad + \frac{\mathbb{P}(Aa|A_2B_1)\mathbb{P}(A_2B_1) + \mathbb{P}(Aa|A_2B_2)\mathbb{P}(A_2B_2)}{\mathbb{P}(AA) + \mathbb{P}(Aa)} \\
&= \frac{1/3(1-p) + 1/6p + 1/3(1-p) + 1/3p}{2/3 - 1/3p + 1/3 + 1/6p} = \frac{4-p}{6-p}.
\end{aligned}$$

例 1.9 一副牌去除大小王 52 张随机的分成 4 堆，每堆 13 张。试计算每堆都有一张 “A” 的概率。

解答：

（乘法公式、全概率公式等可以极大的简化问题）定义如下四个事件：

- $E_1 :=$ {黑桃 A 属于某一堆};
- $E_2 :=$ {黑桃 A 和红桃 A 属于不同堆};
- $E_3 :=$ {黑桃 A, 红桃 A 和梅花 A 属于不同的堆};
- $E_4 :=$ {四张 A 属于不同的堆}.

由题意，目标是要求概率 $\mathbb{P}(E_4)$, 由乘法公式可知

$$\mathbb{P}(E_4) = \mathbb{P}(E_4|E_3)\mathbb{P}(E_3).$$

条件概率 $\mathbb{P}(E_4|E_3) = \frac{13}{49}$, （在给定 E_3 的条件下，事件 E_4 等价于为方片 A 挑选位置），因此只需求 $\mathbb{P}(E_3)$, 同理

$$\mathbb{P}(E_3) = \mathbb{P}(E_3|E_2)\mathbb{P}(E_2) = 26/50\mathbb{P}(E_2).$$

和

$$\mathbb{P}(E_2) = \mathbb{P}(E_2|E_1)\mathbb{P}(E_1) = 39/51.$$

因此

$$\mathbb{P}(E_4) = 13/49 * 13/25 * 13/17.$$

练习三张形状大小完全相同但颜色不同的卡片，第一张两面都是红色，第二张一面红色一面黑色，第三张两面都是黑色。把三张卡片完全遮住，随机抽取一张放在桌上，朝上一面为红色。那么另一面是黑色的概率是多少？

第 2 章 第二章：随机变量

在问题或者要计算的结果之前，如果没有随机变量，第一步是定义相应的随机变量，然后进行运算。特别注意：密度函数和分布函数的表达式，在 $\mathbb{R}$ 上完整的表达式。

2.1 定义和常见分布

1. 随机变量：从样本空间到实数集（数域）的映射。
2. 分布函数：$F(x) := \mathbb{P}(X \leq x)$. 分布函数是定义在实数域取值为 $[0,1]$ 且单调不减的函数，左边极限存在且右连续。分布函数的性质？分布函数的性质：（1）$1 \leq F(x) \leq 1$, 且 $\lim\limits_{x\to-\infty} F(x) = 0$, $\lim\limits_{x\to\infty} F(x) = 1$;（2）$F(x)$ 单调非减：即 $x_1 < x_2$ 时，$F(x_1) \leq F(X_2)$;（3）$F(x)$ 右连续。
3. 若随机变量只有有限个或者可列个取值则称其为离散型随机变量，概率质量函数或频率函数：$\mathbb{P}(X = x_k) = p(x_k) = p_k$.
4. 常见的离散型随机变量
   (a) 单点型分布：$\mathbb{P}(X = c) = 1$.
   (b) 伯努利型随机变量：$\mathbb{P}(X = 0) = p, \mathbb{P}(X = 1) = 1 - p$. 二项分布（随机变量）: $\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$.
   (c) 几何分布-实验首次成功：$\mathbb{P}(X = k) = (1-p)^k p$. 负二项分布：实验第 k 次成功 $\mathbb{P}(X = n) = \binom{n-1}{k-1} p^k (1-p)^{n-k}$.。
   (d) 超几何分布-从罐子中取球：$\mathbb{P}(X = k) = \frac{\binom{r}{k}\binom{n-r}{m-k}}{\binom{n}{m}}$.
   (e) Poisson 分布：$\mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$.
5. **Poisson 定理**
   设 $\lambda > 0$ 为常数，n 为正整数，且满足 $\lim\limits_{n\to\infty} np_n = \lambda$, 则对于任一非负整数 k 有

$$\lim_{n\to\infty} \binom{n}{k} p_n^k (1-p_n)^{n-k} = \frac{\lambda^k}{k!} e^{-\lambda}.$$

6. 连续型随机变量

- 密度函数 $f(x)$: 假如有 $F(x) = \int_{-\infty}^{x} f(y)dy$;
- $\mathbb{P}(X = c) = 0$; （概率为 0 的含义?）
- 分位数（quantile）x_p: 满足 $\int_{-\infty}^{x_p} f(x)dx = p$. 中位数，上下分位数。

7. 连续型随机变量概率密度的性质:

(a) $f(x) \geq 0, -\infty < x < \infty$;

(b) $\int_{-\infty}^{\infty} f(x)dx = 1$; 可以用来确定参数值，如 $f(x) = ce^{-|x|/8}$.

(c) 定义：$F(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^{x} f(t)dt$;

(d) 求概率：$\mathbb{P}(a \leq X \leq b) = \int_{a}^{b} f(x)dx$.

8. 常见的连续分布

- 均匀分布 $U(a, b)$: $f(x) = \begin{cases} \frac{1}{b-a} & x \in (a, b) \\ 0 & Otherwise. \end{cases}$
- 指数分布 $Exp(\lambda)$: $f(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & x \leq 0. \end{cases}$ 指数分布是无记忆的（反过来，如果某非退化连续型随机变量具有无记忆性，那么它服从指数分布）。
- 离散分布中具有无记忆性的是几何分布：直观解释。
- 伽马（Gamma）分布 $\Gamma(r, \lambda)$: $f(x) = \begin{cases} \frac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x} & x > 0 \\ 0 & Otherwise. \end{cases}$ 第 r 个人或者第 r 次冲击到达的时间。
- 正态分布 $N(\mu, \sigma^2)$: $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\{-\frac{(x-\mu)^2}{2\sigma^2}\}$ 当 $\mu = 0, \sigma = 1$ 时，称之为标准正态分布。
- 贝塔（Beta）分布 $\beta(a,)$: $f(x) = \begin{cases} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1} (1-x)^{b-1} & x \in (0, 1) \\ 0 & Otherwise. \end{cases}$

9. 正态分布的性质

- 正态分布的密度函数关于 $x = \mu$ 对称，且在 $x = \mu$ 点取得最大值；
- 保持参数 σ 不变，改变 μ 值，密度函数形状不变，仅仅左右移动；
- 若保持 μ 不变，σ 改变。则 σ 越大越平坦，σ 越小越陡峭；
- 假设 $X \sim N(\mu, \sigma)$, 则 $\frac{X-\mu}{\sigma} \sim N(0, 1)$.

例 2.1 假设随机变量 X 的频率函数为

$$\mathbb{P}(X = x) = c(2/3)^x, x = 1, 2, 3.$$

求 c 的值。

解答：

由频率函数的性质可知

$$c[2/3+(2/3)^2+(2/3)^2]=1$$

因此 $c=\frac{27}{38}$.

例 2.2 假设随机变量 X 服从泊松分布，求 k 使得 $\mathbb{P}(X=k)$ 达到最大。

解答：

因为 X 服从泊松分布，则 $\mathbb{P}(X=k)=\frac{\lambda^k}{k!}e^{-\lambda}$. 则

$$\frac{\mathbb{P}(X=k+1)}{\mathbb{P}(X=k)}=\frac{\lambda}{k+1},$$

这是关于 k 的递减函数，当 $\frac{\lambda}{k+1}>1$ 时，$\mathbb{P}(X=k)$ 是递增的；当 $\frac{\lambda}{k+1}<1$ 时，$\mathbb{P}(X=k)$ 是递减的。因此要使 $\mathbb{P}(X=k)$ 达到最大，当且仅当

$$k=\begin{cases}\lambda \text{ or } \lambda-1, & \lambda\text{为整数}\\ [\lambda] & otherwise.\end{cases}$$

这里 $[\cdot]$ 为取整函数。

例 2.3 设 $0<\mathbb{P}(B)<1$, 试证明事件 A,B 相互独立的充分必要条件是

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

证明：

假如，事件 A,B 相互独立，则 $A,\bar{B}$ 也相互独立，因此

$$\mathbb{P}(A|B)=\frac{\mathbb{P}(AB)}{\mathbb{P}(B)}=\mathbb{P}(A),$$

同理

$$\mathbb{P}(A|\bar{B})=\frac{\mathbb{P}(A\bar{B})}{\mathbb{P}(\bar{B})}=\mathbb{P}(A),$$

因此 $\mathbb{P}(A|B)=\mathbb{P}(A|\bar{B})$.

反之，假设有 $\mathbb{P}(A|B)=\mathbb{P}(A|\bar{B})$，因此

$$\frac{\mathbb{P}(AB)}{\mathbb{P}(B)}=\frac{\mathbb{P}(A\bar{B})}{\mathbb{P}(\bar{B})}=\frac{\mathbb{P}(A)-\mathbb{P}(AB)}{1-\mathbb{P}(B)}.$$

化简可得 $\mathbb{P}(AB)=\mathbb{P}(A)\mathbb{P}(B)$, 即 A,B 相互独立。

结论得证。

例 2.4 假设函数 $F_1(x),F_2(x)$ 分别是随机变量 X_1,X_2 的分布函数，为使得函数 $F(x)=aF_1(x)-bF_2(x)$ 是某一随机变量的分布函数，下列给定数值中应选：

1. $a=3/5,b=-2/5$；

2. $a = 2/3, b = 2/3$;
3. $a = 3/2, b = 1/2$;
4. $a = 1/2, b = -3/2$.

解答:

分布函数需要满足：归一性，单调非减，右连续。只有第一种情况满足所有的性质。

例 2.5 假设随机变量 X_1, X_2 为两个连续型续集变量，他们的概率密度函数分别为 $f_1(x), f_2(x)$, 分布函数分别为 $F_1(x), F_2(x)$, 则:

1. $f_1(x) + f_2(x)$ 必为某一随机变量的密度函数;
2. $f_1(x)f_2(x)$ 必为某一随机变量的密度函数;
3. $F_1(x) + F_2(x)$ 必为某随机变量的分布函数;
4. $F_1(x)F_2(x)$ 必为某随机变量的分布函数。

解答:

需要满足分布函数和密度函数的性质，显然前三种情况都不满足，只有第四种满足。

例 2.6 设随机变量 X 的密度函数为

$$f(x) = \begin{cases} 1/3, & x \in [0,1]; \\ 2/9 & x \in [3,6] \\ 0, & otherwise. \end{cases}$$

如果 k 使得 $\mathbb{P}(X \geq k) = 2/3$, 则 k 的取值为?

解答:

k 的取值为 $(1,3]$ 之间任意值。

例 2.7 假设随机变量 X 的密度函数 $f(x)$ 满足

$$f(x) := \begin{cases} c\sqrt{1-x^2}, & -1 \leq x \leq 1 \\ 0, & otherwise. \end{cases}$$

试求参数 c 的值以及分布函数 $F(x)$.

解答:

由分布函数的性质可知

$$\int_{-\infty}^{\infty} f(x)dx = c\int_{-1}^{1} \sqrt{1-x^2}dx = 1.$$

而由分部积分公式

$$\begin{aligned}\int_{-1}^{1}\sqrt{1-x^2}dx &= x\sqrt{1-x^2}|_{-1}^{1}+\int_{-1}^{1}\frac{x^2}{\sqrt{1-x^2}}dx\\ &= \frac{1}{2}\arcsin x|_{-1}^{1}\\ &= \frac{\pi}{2}.\end{aligned}$$

分布函数为

$$F(x)=\begin{cases}0, & x<0;\\ \frac{x}{\pi}\sqrt{1-x^2}+\frac{1}{\pi}\arcsin x+\frac{1}{2}, & -1\le x<1\\ 1 & 1\le x.\end{cases}$$

练习假设公交车从早上 7 点开始，到达某一站点的时间间隔是 15 分钟，也就说 7:00,7:15，7:30 这类时间到达。如果乘客到达车站的时间服从 7:00-7:30 的均匀分布，那么事件（1）他等公交车的时间不超过 5 分钟；（2）他等公交车的时间超过 10 分钟的概率是多少?

练习设 K 在 $(0,5)$ 上服从均匀分布，求方程

$$4x^2+4Kx+(K+2)=0$$

有实根的概率．

练习设顾客在某银行的窗口等待服务的时间 X (分钟) 服从指数分布，其概率密度为

$$f(x):=\begin{cases}\frac{1}{5}e^{-\frac{x}{5}}, & x\ge 0;\\ 0, & otherwise.\end{cases}$$

某顾客在窗口等待服务，若超过 10 分钟，他就离开．他一个月要到银行 5 次．以 Y 表示一个月内他未等到服务而离开窗口的次数，求 Y 的分布律，并求 $\mathbb{P}(Y\ge 1)$.

例 2.8 假设随机变量 X,Y 均服从正态分布，$X\sim N(\mu,4^2),Y\sim N(\mu,5^2)$, 记 $p_1=\mathbb{P}(X\le\mu-4),p_2=\mathbb{P}(Y\ge\mu+5)$, 则

1. 对任意实数 μ, 都有 $p_1=p_2$；
2. 对任意实数 μ, 都有 $p_1<p_2$；
3. 只对个别实数 μ, 都有 $p_1=p_2$；
4. 对任意实数 μ, 都有 $p_1>p_2$；

解答：

利用正态分布的对称性，和 $\frac{X-\mu}{\sigma}$ 服从标准正态，所以第一个结论正确。

例 2.9 假设随机变量 X 的密度函数为 $f(x)$, 且 $f(x)=f(-x)$, $F(x)$ 是随机变量 X 的分布函数，则对任意实数 a, 下面一定成立的是

1. $F(-a) = 1 - \int_0^a f(x)dx$;
2. $F(-a) = 1/2 - \int_0^a f(x)dx$;
3. $F(-a) = F(a)$;
4. $F(-a) = 2F(a) - 1$;

解答：
利用分布的对称性和分布函数的性质，（排除）可知第二个答案正确。

例 2.10 假设随机变量 $X \sim N(\mu, \sigma^2)$, 则随着参数 σ 的增大，概率 $\mathbb{P}(|X-\mu| \le \sigma)$

1. 单调增大；
2. 单调减小；
3. 保持不变；
4. 非单调变化。

解答：
保持不变

练习假设随机变量 $X \sim N(2, \sigma^2)$, 且 $\mathbb{P}(2 \le X \le 4) = 0.3$, 求 $\mathbb{P}(X \le 0) = ?$

练习假设随机变量 $X \sim N(\mu, \sigma^2)$, 假设方程 $y^2 + 4y + X = 0$ 有实根的概率等于 $\frac{1}{2}$, 则 $\mu = ?$

例 2.11 假设随机变量 X 的密度函数 $f(x) = Ae^{-x^2+x}$, $-\infty < x < \infty$, 试求常数 A 的值。

解答：
由密度函数的性质可知

$$\begin{aligned}\int_{-\infty}^{\infty} f(x)dx &= \int_{-\infty}^{\infty} Ae^{-x^2+x}dx \\ &= \int_{-\infty}^{\infty} Ae^{1/4}e^{-(x-1/2)^2}dx \\ &= Ae^{1/4}\sqrt{\pi} = 1\end{aligned}$$

因此 $A = \frac{1}{\sqrt{\pi}}e^{-1/4}$. 最后一个等式是利用 $e^{-(x-1/2)^2}$ 和 $N(\frac{1}{2}, \frac{1}{\sqrt{2}}^2)$ 的密度函数只差一个常数。

2.2 随机变量的函数

1. 离散随机变量：只需要罗列随机变量函数的取值和相应的概率即可；

2. 连续型随机变量（单调情形）：假设 X 的密度函数为 $f(x)$, 且 $y=g(x)$ 是严格单调函数，其反函数 $h(y)=g^{-1}(y)$ 连续可导函数。则随机变量 $Y=g(X)$ 的密度函数为

$$f_Y(y)=\begin{cases} f(h(y))|h'(y)|, & h(y)\text{有意义} \\ 0 & Otherwise. \end{cases}$$

3. 连续型随机变量 (非单调情形): 假设 X 的密度函数为 $f(x)$, 且 $y=g(x)$ 是在互不相交的区间 $(a_1,b_2),(a_2,b_2),\cdots$ 上逐段严格单调函数，其反函数 $h_1(y),h_2(y),\cdots$ 连续可导函数。则随机变量 $Y=g(X)$ 的密度函数为

$$f_Y(y)=\begin{cases} \sum\limits_i f(h_i(y))|h_i'(y)|, & h_i(y)\text{有意义} \\ 0 & Otherwise. \end{cases}$$

4. 需要掌握的常见情形 $Y=F(X)$, $Y=e^X$, $Y=aX+b$, $Y=X^2$ 等等。

5. 上述需要变换满足特定的条件，而先求分布函数，再求导的方法适用于所有情形。

例 2.12 假设随机变量 X 的密度函数为

$$f(x)=\begin{cases} \frac{2x}{\pi^2}, & 0<x<\pi \\ 0, & otherwise. \end{cases}$$

求 $Y=\sin X$ 的密度函数。

解答：

解法一、因为 $y=\sin x$ 在区域 $(0,\pi/2)$ 和 $(\pi/2,\pi)$ 上的逆函数分别为 $\arcsin y$ 和 $\pi-\arcsin y$, 则 $Y=\sin X$ 的密度函数为

$$\begin{aligned} f_Y(y) &= f(\arcsin y)(\arcsin y)'+f(\pi-\arcsin y)|(\pi-\arcsin y)'| \\ &= \begin{cases} \frac{2}{\pi\sqrt{1-y^2}}, & 0<y<1 \\ 0, & Otherwise. \end{cases} \end{aligned}$$

解法二、由定义 $Y=\sin X$ 的分布函数为

$$F_Y(y)=\mathbb{P}(Y\le y)=\mathbb{P}(X\le \arcsin y)+\mathbb{P}(X\ge \pi-\arcsin y).$$

因此

$$\begin{aligned} f_Y(y) &= F_Y'(y) \\ &= \begin{cases} \frac{2}{\pi\sqrt{1-y^2}}, & 0<y<1 \\ 0, & Otherwise. \end{cases} \end{aligned}$$

例 2.13 **令** X **为服从参数为 2 指数分布的随机变量，试求随机变量** $Y_1 = e^{-2X}$ **和** $Y_2 = 1 - e^{-2X}$ **的分布。**

解答：

解法一、由定义可知

$$
\begin{aligned}
F_{Y_1}(y) &= \mathbb{P}(e^{-2X} \le y) \\
&= \mathbb{P}(X \ge -\frac{\ln y}{2}) \\
&= y
\end{aligned}
$$

因此 $Y_1 \sim U(0,1)$, **同理** $Y_2 \sim U(0,1)$.

第 3 章
第三章：联合分布

3.1 分布函数

1. 二维随机变量的累积分布函数为 $F(x,y)=\mathbb{P}(X\le x, Y\le y)$, 记住分布函数是事件的概率。

2. 二元分布函数的性质：

   (a) 对于固定的 x, 函数 $F(x,y)$ 是关于 y 的右连续函数；对于固定的 y, 函数 $F(x,y)$ 分别是关于 x 的右连续函数。

   (b) 对于固定的 x, 函数 $F(x,y)$ 是关于 y 的单调不减函数；对于固定的 y, 函数 $F(x,y)$ 分别是关于 x 的单调不减函数。

   (c) $0\le F(x,y)\le 1$ 且满足 $F(\infty,\infty)=1, F(-\infty,-\infty)=F(x,-\infty)=F(-\infty,y)=0$.

   (d) 分布函数在矩形 $D:=\{(x,y)|x_1\le x\le x_2, y_1\le y\le y_2\}$ 上的概率 $F(x_2,y_2)-F(x_1,y_2)-F(x_2,y_1)+F(x_1,y_1)\ge 0$.

3. 二维离散随机变量：联合频率函数 $\mathbb{P}(X=x_i, Y=y_j)=p_{ij}$, 对于离散型随机变量，写出分布列很多时候可以极大的简化问题。

4. Remark： 一般来说，对于离散型随机变量，除非特指联合分布函数，让你写出联合分布只需要给出分布列或分布规律就可以。

5. 常用的多维离散分布：多项分布（multinoimal distribution）$p(n_1n_2,\cdots,n_r)=\binom{n}{n_1\cdots n_r}p_1^{n_1}\cdots p_r^{n_r}$ 和多维超几何分布 $p(n_1n_2,\cdots,n_r)=\frac{\binom{N_1}{n_1}\cdots\binom{N_r}{n_r}}{\binom{N}{n}}$.

6. 二维连续型随机变量：联合分布函数 $F(x,y):=\int_{-\infty}^{x}\int_{-\infty}^{y}f(u,v)dudv$，称 $f(x,y)$ 为联合概率密度函数。

7. 联合密度函数 $f(x,y)$: 满足 $f(x,y)\ge 0, \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}f(x,y)dxdy=1$, 且 $\forall D\in\mathbb{R}^2$, $\mathbb{P}((X,Y)\in D)=\iint_D f(x,y)dxdy$. 在 $f(x,y)$ 的连续点处，$\frac{\partial^2 F}{\partial x\partial y}(x,y)=f(x,y)$.（已知分布函数，如何求密度函数）

8. 边际密度和边际分布：$F_X(x)=\int_{-\infty}^{x}\int_{-\infty}^{\infty}f(x,y)dy)dx$, $f_X(x)=\int_{-\infty}^{\infty}f(x,y)dy.$

9. 常用的二维连续型随机变量：二维正态分布

$$\begin{aligned}f(x,y) &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\{-\frac{1}{2(1-\rho^2)}[\frac{(x-\mu_1)^2}{\sigma_1^2}\\ &-2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}]\}.\end{aligned}$$

和多维均匀分布

$$f(x,y)=\begin{cases}\frac{1}{S_D}, & (x,y)\in D;\\ 0, & otherwise.\end{cases}$$

.

10. Remark: 密度函数和分布函数是定义在 $\mathbb{R}$ 或 $\mathbb{R}^2$ 的函数，所以在表示这些函数的时候，需要完整的写出这些函数的取值!!!!!!!

11. Remark: 由高维联合分布可以求出低维边际分布，反之则不一定；不同的联合分布可以有相同的边际分布；多项分布的边际分布仍为边际分布或二项分布；二维正态的边际分布是一维正态分布。

例题解析

例3.1 假设二维随机变量 (X,Y) 的联合密度函数为

$$f(x,y)=\begin{cases}4xy, & 0\le x\le 1, 0\le y\le 1;\\ 0, & otherwise.\end{cases}$$

1. $\mathbb{P}(0<X<0.5, 0.25<Y<1)$;
2. $\mathbb{P}(X=Y)$;
3. $\mathbb{P}(X\le Y)$;
4. (X,Y) 的联合分布函数。

解答：

1、由密度函数的性质可知

$$\begin{aligned}\mathbb{P}(0<X<0.5,0.25<Y<1) &= \int_0^{0.5}\int_{0.25}^{1}4xydydx\\ &= \int_0^{0.5}15/8xdx\\ &= 15/16x^2|_0^{0.5}=15/64\end{aligned}$$

2、因为连续型随机变量在单点的概率等于 0，因此可只有 $\mathbb{P}(X=Y)=0$.

3、事件 $\{X \le Y\}$ 对应的 $\mathbb{R}^2$ 上的点集为 $D := \{(x,y) : 0 \le x \le 1, 0 \le y \le 1, x \le y\}$. 因此

$$\begin{aligned}\mathbb{P}(X \le Y) &= \iint\limits_D 4xydydx \\ &= \int_0^1 \int_0^y 4xydxdy \\ &= \frac{x^2}{2}|_0^1 = 0.5\end{aligned}$$

4、当 $x < 0$ 或 $y < 0$ 时，因为 $\{X < 0\}, \{Y < 0\}$ 是不可能事件，因此此时 $F(x,y) = 0$. 当 $x \ge 1$ 且 $y \ge 1$ 时，因为 $\{X \ge 1, Y \ge 1\}$ 是不可能事件，因此此时 $F(x,y) = 1$. 当 $0 \le x < 1, 0 \le y < 1$ 时，由定义

$$F(x,y) = \int_0^x \int_0^y 4uvdudv = x^2y^2.$$

更进一步，当 $0 \le x < 1, y \ge 1$ 时，由定义

$$\begin{aligned}F(x,y)) &= \int_0^x \int_0^y 4uvdudv \\ &= \int_0^x \int_0^1 4uvdudv \\ &= x^2.\end{aligned}$$

同理，当 $x \ge 1, 0 \le y < 1$, $F(x,y) = y^2$. 由此可知

$$F(x,y) = \begin{cases} 0, & x < 0 \, or \, y < 0; \\ x^2y^2, & 0 \le x < 1, 0 \le y < 1; \\ x^2, & 0 \le x < 1, y \ge 1; \\ y^2, & x \ge 1, 0 \le y < 1; \\ 1, & x \ge 1, y \ge 1. \end{cases}$$

例3.2 假设二维随机变量 (X,Y) 的分布函数为 $F(x,y)$, 则随机变量 (Y,X) 的分布函数 $F_1(x,y) =?$

解答:

由定义可知

$$\begin{aligned}F_1(x,y) &= \mathbb{P}(Y \le x, X \le y) \\ &= \mathbb{P}(X \le y, Y \le x) = F(y,x).\end{aligned}$$

例3.3 假设随机变量 X_1, X_2 的分布列如下表，且满足 $\mathbb{P}(X_1X_2 = 0) = 1$, 试求概率 $\mathbb{P}(X_1 = X_2) =?$

X_i	-1	0	1
p	0.25	0.5	0.25

解答：

(显然直接去算会很复杂，但假如写出分布列)

X_1 \ X_2	-1	0	1
-1	0	p_{12}	0
0	p_{21}	p_{22}	p_{23}
1	0	p_{32}	0

利用边际分布列和联合分布列之间的关系可知 $p_{12} = p_{21} = p_{23} = p_{32} = 0.25$, 因此 $p_{22} = 0$, 从而 $\mathbb{P}(X_1 = X_2) = 0$.

例3.4 假设随机变量 (X, Y) 的联合分布函数为

$$F(x, y) = A(B + \arctan x/2)(C + \arctan y/2),$$

其中 A, B, C 为常数。

- 求参数 A, B, C 的值；
- (X, Y) 的联合密度函数为?
- 求随机变量 X 和 Y 的边际分布函数和密度函数，
- 求 $\mathbb{P}(X > 2)$.

解答：

1、由分布函数的性质可知 $F(-\infty, y) = F(x, -\infty) = 0$, 可只有 $A = B = \pi/2$. 同理，$F(\infty, \infty) = 1$ 可只有 $A = \frac{1}{\pi^2}$.

2、由第一问 $F(x, y) = \frac{1}{\pi^2}(\pi/2 + \arctan x/2)(\pi/2 + \arctan y/2)$, 因此联合密度函数

$$\begin{aligned} f(x, y)) &= \frac{\partial^2 F}{\partial x \partial y}(x, y) \\ &= \frac{1}{\pi^2} \frac{4}{(4 + x^2)(4 + y^2)}. \end{aligned}$$

3、X 的边际分布函数 $F_X(x) = F(x, \infty) = \frac{1}{\pi}(\pi/2 + \arctan x/2)$, 因此边际密度

$$f_X(x) = F'_X(x) = \frac{4}{\pi(4 + x^2)}.$$

同理，Y 的边际分布函数和密度函数为 $F_Y(y) = \frac{1}{\pi}(\pi/2 + \arctan y/2)$ 和 $f_Y(y) = \frac{4}{\pi(4+y^2)}$.

4、由之前结论

$$\mathbb{P}(X > 2) = 1 - F_X(2) = 1/4.$$

例3.5 （蒲丰投针问题）平面上有一些平行线，他们之间的距离是 D, 一根长为 L 针随机的投在平面上，其中 $D \geq L$. 证明针正好与一条直线相交的概率为 $2L/\pi D$.

解答：

确定针在平面上的位置至少需要两个变量，这里选取针的中点距离 X 和针与平行线之间的夹角 Y. 因为是随机投针，所以 $X \sim U(0, D/2), Y \sim U(0, \pi/2)$. 随机变量 (X, Y) 的联合密度函数为

$$f(x,y) = \begin{cases} \frac{4}{\pi D}, & 0 < x < D/2, 0 < y < \pi/2; \\ 0, & otherwise. \end{cases}$$

针与平行线相交可以等价的表述为 $\{X \leq L/2 \sin Y\}$, 对应的点集 $D := \{(x, y) : 0 < x < D/2, 0 < y < \pi/2, x \leq L/2 \sin y\}$. 则针与平行线相交的概率为

$$\mathbb{P}(X \leq L/2 \sin Y) = \int_0^{\pi/2} \int_0^{L/2 \sin y} f(x,y) dx dy = 2L/\pi D.$$

例3.6 在椭圆内部随机的选择一个点，椭圆方程为

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

计算坐标 x 和 y 的边际密度。

解答：

由题意 (X, Y) 的联合密度函数为

$$f(x,y) = \begin{cases} \frac{1}{\pi ab}, & -a < x < a, -b < y < b; \\ 0, & otherwise. \end{cases}$$

(求边际密度，就相当于把该随机变量对应变量的看做常数，对另一个变量在积分区域内求积分)，对随机变量 X, 其边际密度为

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f(x,y) dy \\ &= \int_{-b\sqrt{1-x^2/a^2}}^{b\sqrt{1-x^2/a^2}} \frac{1}{\pi ab} dy \\ &= \frac{2\sqrt{1-x^2/a^2}}{\pi a}. \end{aligned}$$

因此

$$f_X(x)) = \begin{cases} frac2\sqrt{1-x^2/a^2}\pi a, & -a < x < a; \\ 0, & otherwise. \end{cases}$$

同理，随机变量 Y 的边际密度为

$$f_Y(y)) = \begin{cases} \frac{2\sqrt{1-y^2/b^2}}{\pi b}, & -b < y < b; \\ 0, & otherwise. \end{cases}$$

例3.7 假设二维随机变量 (X, Y) 的联合密度函数为

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

试求常数 c 和 $\mathbb{P}(X>Y)$.

解答：

由密度函数的性质可知

$$\begin{aligned} \iint_{\mathbb{R}^2} f(x,y)dxdy &= c\int_{-1}^{1}\int_{x^2}^{1} x^2ydydx \\ &= 4c/21 \\ &= 1. \end{aligned}$$

由此可知 $c=21/4$, 类似的，事件 $\{X>Y\}$ 对应的 $\mathbb{R}^2$ 上的点集为 $D:=\{(x,y): x^2< y\le 1, x>y\}$. 因此

$$\begin{aligned} \mathbb{P}(X\,Y) &= \iint\limits_{D} 21/4x^2ydydx \\ &= \int_0^1\int_{x^2}^{x} 21/4x^2ydydx \\ &= 3/20 \end{aligned}$$

练习 3.1 试求一下两个联合密度函数的边际密度函数

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

和

$$g(x,y)=\begin{cases} (0.5+x)(0.5+y), & 0\le x\le 1, 0\le y\le 1;\\ 0, & otherwise. \end{cases}$$

练习 3.2 假设随机变量 A, B, C 为独立随机变量，且均服从 $(0,1)$ 上的均匀分布。

- 求 A, B, C 的联合分布函数；
- 求方程 $Ax^2+Bx+C=0$ 所有的根都是实数的概率。

练习 3.3 证明函数

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

为一联合密度函数，并求相应的边际分布。

练习 3.4 一个射手进行射击，击中目标的概率为 p, 射击击中目标两次后停止。以 X 表示首次击中目标的射击次数，Y 表示总共进行的射击次数，试求 X, Y 的联合分布列和边际分布列。

练习 3.5 从 $(0,1)$ 中任取两个数，求其积不小于 $3/16$ 且其和不大于 1 的概率?

3.2 独立性和条件分布

1. 对于随机变量 X,Y, 有 $F(x,y)=F_X(x)F_Y(y)$, 则称 X,Y 相互独立。具体来说：离散随机变量 $\mathbb{P}(X=x_i,Y=y_j)=\mathbb{P}(X=x_i)\mathbb{P}(Y=y_j)$ 和连续型随机变量 $f(x,y)=f_X(x)f_Y(y)$.

2. 二维正态随机变量 X,Y 相互独立等价于参数 $\rho=0$.

3. 两个（连续型）随机变量相互独立当且仅当 $f(x,y)=h(x)g(y)$.

4. Remark: 三个随机变量两两相互独立并不一定独立，反例：假如联合密度函数形如
$$f(x,y,z)=\phi(x)\phi(y)\phi(z)(1+\sin x\sin y\sin z).$$
其中 $\phi(\cdot)$ 为标准正态的密度函数。

5. 二维离散型随机变量：条件频率函数 $\mathbb{P}(X=x_i|Y=y_j)=\frac{\mathbb{P}(X=x_i,Y=y_j)}{\mathbb{P}(Y=y_j)}$.

6. 二维连续型随机变量：条件密度 $f_{X|Y}(x|y)=\frac{f(x,y)}{f_Y(y)}$ 和条件分布 $F_{X|Y}(x|y)=\int_{-\infty}^{x}f_{X|Y}(u|y)du$.

7. 连续型随机变量的 Bayes 法则：$f(x,y)=f_{Y|X}(y|x)f_X(x)$ 以及连续型随机变量的全概率公式 $f_Y(y)=\int_{-\infty}^{\infty}f_{Y|X}(y|x)f_X(x)dx$.

8. Remark: 边际分布只和 X 或 Y 本身有关，和其他随机变量无关；条件分布 $f_{Y|X}(y|x)$ 大多数情况下在函数形式和取值范围依赖于条件 $X=x$, 思考什么时候条件分布与 x 无关。

例3.8 在 $\triangle ABC$ 内任取一点 P, 在底边 BC 上任取一点 Q, 求直线 PQ 与线段 AB 相交的概率。

解答：

不妨假设 B 为原点，BC 位于 x 正半轴上，长度为 l. 则解法一、（微分法）由题意 Q 点落在小区间 $(x,x+dx)$, 要使得直线 PQ 与线段 AB 相交当且仅当 P 落在三角形 ABP 内。因此概率为 $\frac{S_{\triangle ABP}}{S_{\triangle ABC}}dx/l$, 则直线 PQ 与线段 AB 相交的概率为
$$\int_0^l\frac{S_{\triangle ABP}}{S_{\triangle ABC}}dx/l=\int_0^l\frac{x}{l}dx/l=1/2.$$
解法二、假设随机变量 X 表示 Q 在 x 轴上的位置，Y 表示 P 点在平面 $\mathbb{R}^2$ 上的位置。则可知 (X,Y) 的联合密度函数为
$$f(x,y)=\begin{cases}\frac{1}{lS_{\triangle ABC}}, & 0<x<l, y\in\triangle ABC;\\ 0, & otherwise.\end{cases}$$

记点集 D 表示事件直线 PQ 与线段 AB 相交，则概率为

$$\iiint_D \frac{S_1}{lS_{\triangle ABC}}dxdy = \int_0^l \iint_{S_{\triangle ABP}} \frac{S_1}{lS_{\triangle ABC}}dydx = 1/2.$$

练习 3.6 在一个圆周上随机的取 n 个点，求它们恰好落在同一个半圆周上的概率。令 $P_1, P_2, \cdots, P_n$ 表示这 n 个点，令事件 A 表示事件“所有的点恰好落在某个半圆上”，令事件 A_i 表示事件” 所有的点都落在以 P_i 为起点，顺时针方向的一个半圆周上。

- 将事件 A 用事件 $A_i, i = 1, 2, \cdots, n$ 表示出来；
- A_i 之间是不是互不相容；
- 求 $\mathbb{P}(A)$.

例3.9 假设二维随机变量 (X, Y) 的分布列如下表，试问当 α, β 取什么值时，X 和 Y 是相互独立的?

X \ Y	1	2	3
1	1/6	1/9	1/18
2	1/3	α	β

解答：

写出边际分布列，由独立性可得

X \ Y	1	2	3	$p_{i\cdot}$
1	1/6	1/9	1/18	1/3
2	1/3	α	β	2/3
$p_{\cdot j}$	1/2	$1/9+\alpha$	$1/18+\beta$	1

$\mathbb{P}(X = 1, Y = 2) = 1/3 * (1/9 + \alpha) = 1/9$ 以及 $\mathbb{P}(X = 1, Y = 3) = 1/3 * (1/18 + \beta) = 1/18$.

练习 3.7 假设二维随机变量 (X, Y) 的分布列如下表，试问当 a, b, c 取什么值时，X 和 Y 是相互独立的?

X \ Y	y_1	y_2	y_3	$p_{\cdot j}$
x_1	a	1/9	c	
x_2	1/9	b	1/3	
$p_{i\cdot}$				

例3.10 假设二维随机变量 (X, Y) 的联合密度函数

$$f(x, y) = \begin{cases} 6e^{-2x}e^{-3y}, & 0 < x < \infty, 0 < y < \infty; \\ 0, & otherwise. \end{cases}$$

试判断 X, Y 是否独立？反之，假如

$$f(x,y)=\begin{cases}24xy, & 0<x<1, 0<y<1, 0<x+y<1;\\ 0, & otherwise.\end{cases}$$

X, Y 是否独立？

解答：

对于第一种情况，显然 x, y 的支集（使得函数非 0 的区域）相互独立，且 $f(x,y)=h(x)g(y)$ 可以判断 X, Y 独立。对于第二种情况，x, y 的定义域是相关的，由定义

$$f_X(x)=\int_{-\infty}^{\infty}f(x,y)dy=\int_0^{1-x}24xydy=12x(1-x)^2.$$

即 X 的边际密度为

$$f_X(x)=\begin{cases}12x(1-x)^2, & 0<x<1;\\ 0, & otherwise.\end{cases}$$

同理 Y 的边际密度为

$$f_Y(y)=\begin{cases}12y(1-y)^2, & 0<x<1;\\ 0, & otherwise.\end{cases}$$

可以判断 X, Y 不独立。

例3.11 令 X, Y, Z 服从 $(0,1)$ 上的均匀分布，试求 $\mathbb{P}(X>YZ)$.

解答：

（此题不能画图找积分限）事件 $\{X>YZ\}$ 对应的点集为 $D:=\{(x,y,z):0<x<1, 0<y<1, 0<z<1, x>yz\}$. 由定义可知

$$\begin{aligned}\mathbb{P}(X>YZ) &= \iiint\limits_D dxdydz\\ &= \int_0^1\int_0^1\int_{yz}^1 dxdydz\\ &= 3/4\end{aligned}$$

例3,12 假设某班车上起始站上车的人数 X 服从参数为 $\lambda(\lambda>0)$ 的 Poisson 分布。每位乘客在中途下车的概率为 $p(0<p<1)$, 且他们在中途下车与否是相互独立的。用 Y 表示在中途下车的人数，求

1. 在发车时有 n 个人上车的条件下，有 m 个人在中途下车的概率；
2. 二维随机变量 (X,Y) 的联合分布。

解答：

1、由题意每个人下车与否相互独立，且概率都为 p, 则在给定 $X=n$ 的条件下，中途下车的人 $Y\sim B(n,p)$. 因此 $\mathbb{P}(Y=m|X=n)=\binom{n}{m}p^m(1-p)^{n-m}$.
2、由条件分布的定义可知有

$$\mathbb{P}(X=n,Y=m)=\frac{\lambda^n}{n!}e^{-\lambda}\binom{n}{m}p^m(1-p)^{n-m},\ n=0,1,2,\cdots,m=0,1,\cdots,n,$$

例 **3.13** 设二维随机变量 (X,Y) 的联合概率密度为

$$f(x,y)=\begin{cases}kx, & 0<x<1,0<y<x;\\ 0, & otherwise.\end{cases}$$

1. 确定常数 k;
2. 求条件概率密度 $f_{X|Y}(x|y)$;
3. 试求概率 $\mathbb{P}(X<1/2|Y=1/2)$.

解答：
1、由密度函数的性质可知

$$\iint_{\mathbb{R}^2}f(x,y)dxdy=\int_0^1\int_0^x kxdydx=1.$$

因此 $k=3$.
2、Y 的边际函数

$$f_Y(y)=\int_{-\infty}^{\infty}f(x,y)dx=\int_y^1 3xdx=\frac{3}{2}(1-y^2),y\in(0,1).$$

因此条件密度函数为

$$f_{X|Y}(x|y)=\frac{f(x,y)}{f_Y(y)}=\begin{cases}\frac{2x}{1-y^2}, & y<x<1;\\ 0, & otherwise.\end{cases}$$

3、由第二问可知条件密度函数当 $x<y$ 时 $f_{X|Y}(x|y)=0$，因此 $\mathbb{P}(X<1/2|Y=1/2)=0$.
Remark: 记住求密度函数和分布函数，要完整的写出在整个定义域上的取值，不然直接用第二问得到的函数求积分得到的值为 $1/3$, 虽然费力，但肯定做错了!!!!!

例 **3.14** 假设二维随机变量 (X,Y) 的联合密度函数

$$f(x,y)=\begin{cases}\frac{e^{-x/y}e^{-y}}{y}, & 0<x<\infty,0<y<\infty;\\ 0, & otherwise.\end{cases}$$

试求 $\mathbb{P}(X>1|Y=y)$.

解答：

解法一、由题意，随机变量 Y 的边际密度为

$$f_Y(y) = \int_0^\infty f(x,y)dx = e^{-y}, x > 0.$$

则条件密度 $f_{X|Y}(x|y)$ 为

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)} = \begin{cases} \frac{e^{-x/y}}{y}, & 0 < x < \infty; \\ 0, & otherwise. \end{cases}$$

解法二、由条件密度定义

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)} = \begin{cases} c(y)e^{-x/y}, & 0 < x < \infty; \\ 0, & otherwise. \end{cases}$$

这里 $c(y)$ 是依赖于 y 的常数，然后利用密度函数性质 $\int_0^\infty f_{X|Y}(x|y)dx = 1$, 可知 $c(y) = 1/y$. （这是求条件概率经常用的方法，但要特别注意条件密度的积分限）剩下的求概率可以简单的对条件密度积分。

例3.15 进行 $n+m$ 次独立重复试验，假设每次成功的概率相同。然而，假设这个概率不是固定的，而是一个随机变量，其分布是 $(0,1)$ 上的均匀分布。试问已知 $n+m$ 次试验中有 n 次成功的条件下每次试验成功的概率?

解答：

假设随机变量 X 表示 $n+m$ 次试验中成功的次数，Y 表示成功的概率。由题意可知

$$Y \sim U(0,1), \;\; X|Y = p \sim B(n+m, p).$$

因此联合密度函数

$$f(n,p) = \begin{cases} \binom{n+m}{n} p^n (1-p)^m, & n = 1, 2, \cdots, 0 < p < 1; \\ 0, & otherwise. \end{cases}$$

于是 $f_{Y|X}(p|n) = \frac{f(n,p)}{\mathbb{P}(X=n)} = c(n)p^n(1-p)^m$, 这里 $c(n)$ 是依赖于 n 的常数。仔细观察 $f_{Y|X}(p|n)$ 的形式和 β 分布函数一样，因此 $Y|\{X = n\}$ 服从 β 分布。

练习 3.8 已知随机变量 Y 的密度函数为

$$f_Y(y) = \begin{cases} 5y^4, & 0 < y < 1; \\ 0, & otherwise. \end{cases}$$

在给定 $Y = y$ 的条件下，随机变量 X 的条件密度为

$$f_{X|Y}(x|y) = \begin{cases} \frac{3x^2}{y^3}, & 0 < x < y < 1; \\ 0, & otherwise. \end{cases}$$

求概率 $\mathbb{P}(X > 1/2)$.

练习 **3.9** 假设 X, Y 的联合密度函数为

$$f(x,y)=\begin{cases}\frac{6}{7}(x^2+\frac{xy}{2}, & 0<x<1, 0<y<2;\\ 0, & otherwise.\end{cases}$$

为一联合密度函数，并求相应的边际分布。

- 试证明 $f(x,y)$ 是一个联合密度函数；
- 求 X 的边际密度函数；
- 求 $\mathbb{P}(X>Y)$;
- 求 $\mathbb{P}(Y>1/2|X<1/2)$.

练习 **3.10** 证明函数

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

为一联合密度函数，并求相应的边际分布。

例**3.16** 设随机变量 X_1, X_2, X_3, X_4, X_5 为独立同分布的连续型随机变量，其公共分布函数和密度函数分别为 F 和 f. 试求概率 $\mathbb{P}(X_1<X_2<X_3<X_4<X_5)$.

解答：

事件 $\{X_1<X_2<X_3<X_4<X_5\}$ 对应的点集为

$$\{(x_1,x_2,x_3,x_4,x_5):x_1<x_2<x_3<x_4<x_5\}.$$

因此

$$\begin{aligned}\mathbb{P}(X_1<X_2<X_3<X_4<X_5) &= \int\cdots\int f(x_1)f(x_2)f(x_3)f(x_4)f(x_5)dx_1\cdots dx_5\\ &= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{x_3}\int_{-\infty}^{x_2} f(x_1)f(x_2)f(x_3)f(x_4)f(x_5)dx_1\cdots dx_5\\ &= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{x_2} F(x_2)f(x_2)f(x_3)f(x_4)f(x_5)dx_2\cdots dx_5\\ &= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{x_3} \frac{F^2(x_3)}{2}f(x_3)f(x_4)f(x_5)dx_3dx_4dx_5\\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{x_4} \frac{F^3(x_4)}{6}f(x_4)f(x_5)dx_4dx_5\\ &= \int_{-\infty}^{\infty} \frac{F^4(x_5)}{4!}f(x_5)dx_5=1/5!\end{aligned}$$

练习 **3.11** 设二维随机变量 (X,Y) 的联合概率密度为

$$f(x,y)=\begin{cases}24(1-x)y, & 0<y<x<1;\\ 0, & otherwise.\end{cases}$$

试求 $0<y<1$ 时，求条件密度函数 $f_{X|Y}(x|y)$.

例3.17 假设 X,Y 具有联合密度函数

$$f(x,y)=c\sqrt{1-x^2-y^2},\ x^2+y^2\le 1.$$

1. 计算 c;

2. 画出联合密度函数图形；

3. 计算概率 $\mathbb{P}(X^2+Y^2\le 1/2)$;

4. 计算 X 和 Y 的边际密度，并判断是否独立；

5. 计算条件密度。

解答：

1、由密度函数性质 $c\iint\limits_{x^2+y^2\le 1}\sqrt{1-x^2-y^2}dxdy=1$, 因此只需计算积分 $\iint\limits_{x^2+y^2\le 1}\sqrt{1-x^2-y^2}dxdy$ 即可。有下面三种方法：方法一、假如你可以看出 $\sqrt{1-x^2-y^2}$ 是上半球面，因此 $\iint\limits_{x^2+y^2\le 1}\sqrt{1-x^2-y^2}dxdy$ 等于半球体积，利用球的体积公式 $V=\frac{4\pi}{3}R^3$, 可知 $c=\frac{3}{2\pi}$;
方法二、利用极坐标变换 $x=r\sin\theta,y=r\cos\theta$, 可知

$$\iint\limits_{x^2+y^2\le 1}\sqrt{1-x^2-y^2}dxdy=\int_0^{2\pi}\int_0^1 r\sqrt{1-r^2}drd\theta=\frac{2\pi}{3}.$$

方法三、化成累次积分

$$\iint\limits_{x^2+y^2\le 1}\sqrt{1-x^2-y^2}dxdy=\int_{-1}^1\int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}}\sqrt{1-x^2-y^2}dxdy=\frac{2\pi}{3}.$$

这里省略了中间步骤，因为第三步需要求边际密度。2、省略；3、利用极坐标变换；4、由边际密度定义，可知 x 的边际密度为

$$\begin{aligned}
f_Y(y) &= \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}}\sqrt{1-x^2-y^2}dx\ (\text{令}l=\sqrt{1-y^2}\\
&= \int_{-l}^{l}\sqrt{l^2-x^2}dx\ (u=x/l)\\
&= l^2\int_{-1}^{1}\sqrt{1-u^2}du.
\end{aligned}$$

因此只需计算 $\int_{-1}^1\sqrt{1-u^2}du$ 即可，如果你看出这是半圆的面积，直接带圆的面积公式即可，否则

分步积分法

$$\begin{aligned}\int_{-1}^{1}\sqrt{1-x^2}dx &= x\sqrt{1-x^2}|_{-1}^{1}+\int_{-1}^{1}\frac{x^2}{\sqrt{1-x^2}}dx,\ (u=\sqrt{1-x62},v=x)\\ &= \int_{-1}^{1}\frac{x^2-1+1}{\sqrt{1-x^2}}dx\\ &= \int_{-1}^{1}\sqrt{1-x^2}dx+\arcsin x|_{-1}^{1}\end{aligned}$$

因此 $\int_{-1}^{1}\sqrt{1-u^2}du=\pi/2$.

三角换元法

$$\begin{aligned}\int_{-1}^{1}\sqrt{1-x^2}dx &= \int_{-\pi/2}^{\pi/2}\sqrt{1-\sin^2\theta}d\sin\theta,\ (x=\sin\theta)\\ &= \int_{-\pi/2}^{\pi/2}\cos^2\theta d\theta\\ &= \int_{-\pi/2}^{\pi/2}\frac{\cos 2\theta+1}{2}d\theta=\pi/2\end{aligned}$$

5、略。

例3.18 将长度为 d 的木棍任意截取一段，再将剩下的木棒分为两段，求三段木棍可以做成三角形的概率。

解答：

假设第一段的长度为 X, 第二段的长度为 Y. 则由题意可知 $X\sim U(0,d), Y|X\sim U(0,d-X)$ 则 (X,Y) 的联合密度函数为

$$f(x,y)=\begin{cases}\frac{1}{d(d-x)}, & 0<x+y<d;\\ 0, & otherwise.\end{cases}$$

三段木棍 $X,Y,d-X-Y$ 做成三角形当且仅当

$$\begin{cases}X+Y>d-X-Y;\\ X+(d-X-Y)>Y;\\ Y+(d-X-Y)>X.\end{cases}$$

化简上面不等式可知时间对应的点集为

$$D:=\{(x,y):x<d/2,y<d/2,d/2<x+y<d\}$$

因此概率为

$$\begin{aligned}\iint_D f(x,y)dxdy &= \int_0^d\int_{d/2-x}^{d/2}\frac{1}{d-x}dydx\\ &= \int_0^d\frac{x}{d-x}dydx\\ &= \ln 2-1/2\end{aligned}$$

3.3 随机变量的函数

1. 基本问题：如果二维随机变量 (X,Y) 的联合密度函数为 $f(x,y)$, 则 $Z=g(X,Y)$ 服从什么分布?

2. 随机变量和 $Z=X+Y$ 的概率密度为 $f_Z(z)=\int_{-\infty}^{\infty}f(z-y,y)dy=\int_{-\infty}^{\infty}f(x,z-x)dx$. 如果 X 和 Y 是相互独立的，那么随机变量 $Z=X+Y$ 的概率密度为 $f_Z(z)=\int_{-\infty}^{\infty}f_X(z-y)f_Y(y)dy=\int_{-\infty}^{\infty}f_Y(z-x)f_X(x)dx$. (卷积公式，convolution)

3. 常见连续型随机变量的可加性：正态分布：如果 $X_1,\cdots,X_n$ 相互独立且满足 $X_i\sim N(\mu_i,\sigma_i^2)$ 则对于不全为 0 的常数 $a_1,a_2,\cdots,a_n$ 随机变量

$$a_1X_1+a_2X_2+\cdots+a_nX_n\sim N(\sum_{i=1}^n a_i\mu_i,\sum_{i=1}^n a_i^2\sigma_i^2).$$

指数分布：如果 $X_1,\cdots,X_n$ 相互独立且满足 $X_i\sim Exp(\lambda)$ 则随机变量

$$X_1+X_2+\cdots+X_n\sim\Gamma(n,\lambda).$$

4. 离散随机变量 $Z=X+Y$ 的频率为 $\mathbb{P}(Z=k)=\sum_{i=1}^{k-1}\mathbb{P}(X=i)\mathbb{P}(Y=k-i)=\sum_{i=1}^{k-1}\mathbb{P}(Y=i)\mathbb{P}(X=k-i)$. (离散卷积公式，convolution)

5. Poisson 分布：如果 $X_1,\cdots,X_n$ 相互独立且满足 $X_i\sim P(\lambda_i)$ 则随机变量

$$X_1+X_2+\cdots+X_n\sim P(\sum_{i=1}^n\lambda_i).$$

6. 连续型随机变量积 $Z=XY$ 的密度函数为 $f_Z(z)=\int_{-\infty}^{\infty}f(v,z/v)\frac{1}{|v|}dv=\int_{-\infty}^{\infty}f(z/v,v)\frac{1}{|v|}dv$.

7. 连续型随机变量商 $Z=X/Y$ 的密度函数为 $f_Z(z)=\int_{-\infty}^{\infty}f(v,v/z)\frac{|v|}{z^2}dv=\int_{-\infty}^{\infty}f(zv,v)|v|dv$.

8. 假设连续可微分函数

$$\begin{cases}x=x(u,v)\\ y=y(u,v).\end{cases}$$

将 xy 上有界闭区域 D 一一映射到 uv 平面上 D，其 Jacobe 行列式为

$$J=\frac{\partial(x,y)}{\partial(u,v)}=\begin{vmatrix}\frac{\partial x}{\partial u}&\frac{\partial x}{\partial v}\\ \frac{\partial y}{\partial u}&\frac{\partial y}{\partial v}\end{vmatrix}$$

则 $\iint\limits_D f(x,y)dxdy=\iint\limits_{D'}f(x(u,v),y(u,v))|J|dudv$.

9. n 个独立同分布随机变量极值统计量 $X_{(n)}, X_{(1)}$ 的密度函数分别为 $f_{\max}(z) = nf(z)(F(z))^{n-1}, f_{\min}(z) = nf(z)(1-F(z))^{n-1}$. 第 k 个次序统计量的密度函数 $f_k(x) = \frac{n!}{(k-1)!(n-k)!}f(x)(F(x))^{k-1}(1-F(x))^{n-k}$.

10. $X_{(k)}, X_{(j)}$ 的联合分布：$f_{kj}(x,y) = \frac{n!}{(k-1)!(j-k-1)!(n-j)!}f(x)f(y)(F(x))^{k-1}(F(y)-F(x))^{j-k-1}(1-F(y))^{n-j}$.

例 3.19 令 U_1 和 U_2 是 $[0,1]$ 上相互独立的均匀随机变量，计算并画出 $S = U_1 + U_2$ 的密度函数。

解答：

解法一、直接法：由定义，$S = U_1 + U_2$ 的分布函数

$$
\begin{aligned}
F(S \le z) &= \mathbb{P}(U_1 + U_2 \le z) \\
&= \begin{cases} \int_0^z \int_0^{z-u_1} 1du_2du_1, & z \le 1; \\ \int_0^{z-1}\int_0^1 1du_2du_1 + \int_{z-1}^1\int_0^{z-u_1} 1du_2du_1, & z \ge 1. \end{cases}
\end{aligned}
$$

因此 S 的密度函数为

$$
f(z) = \begin{cases} z, & 0 < z < 1; \\ 2-z, & 1 < z < 2; \\ 0, & otherwise. \end{cases}
$$

解法二、利用卷积公式 $S = U_1 + U_2$ 的密度函数为 =

$$
\begin{aligned}
f_S(z) &= \int_{-\infty}^{\infty} f_{U_1}(z-y)f_{U_2}(y)dy \\
&= \begin{cases} \int_0^z 1dy, & z \le 1; \\ \int_{z-1}^1 1dy, & z \ge 1. \end{cases} \\
&= \begin{cases} z, & z \le 1; \\ 2-z, & z \ge 1. \end{cases}
\end{aligned}
$$

Remark：注意上面卷积公式的积分区域，要保证 $f_{U_1}(z-y)f_{U_2}(y) \neq 0$, 需要满足 $0 < y < 1, 0 < z-y < 1$, 然后分情况讨论。

例 3.20 计算两个独立均匀随机变量商的密度。

解答：

不妨假设 $X \sim U(0,a), Y \sim U(0,b)$.

解法一、问题等价于求事件 $\{X/Y \le z\}$ 的概率，如果 Y 不变好，问题等价于求事件 $\{X \le zY\}$ 的概率，可以根据与概率密度支集围成的图形的面积即可。解法二、连续型随机变量商 $Z = X/Y$ 的密度函数为 $f_Z(z) = \int_{-\infty}^{\infty} f(zv,v)|v|dv$, 因为 $X \sim U(0,a), Y \sim U(0,b)$, 所以 $f(zv,v) \neq 0$ 当且仅当 $0 < zv < a, 0 < v < b$. 分下面几种情况进行讨论：

1、当 $z < 0$, 事件是不可能事件，此时 $f_Z(z) = 0$;

2、当 $z>0$, 如果 $z<a/b$, 则积分区域为 $0<v<b$ 此时 $f_Z(z)=\int_0^b f(zv,v)|v|dv=\frac{b}{2a}$;
3、当 $z>0$, 如果 $z>a/b$, 则积分区域为 $0<v<\frac{a}{z}$ 此时 $f_Z(z)=\int_0^{a/z} f(zv,v)|v|dv=\frac{a}{2bz^2}$.
所以，$Z=X/Y$ 的密度函数为

$$f(z)=\begin{cases}\frac{b}{2a}, & 0<z<a/b;\\ \frac{a}{2bz^2}, & a/b<z<\infty;\\ 0, & otherwise.\end{cases}$$

对于更一般的情况，假设 $X\sim U(a,b), Y\sim U(c,d)$.
连续型随机变量商 $Z=X/Y$ 的密度函数为 $f_Z(z)=\int_{-\infty}^{\infty} f(zv,v)|v|dv$, 因为 $X\sim U(a,b), Y\sim U(c,d)$, 所以 $f(zv,v)\neq 0$ 当且仅当 $a<zv<b, c<v<d$. 分下面几种情况进行讨论：
1、当 $z<0, z<a/c\ or\ z>b/d$, 此时事件是不可能事件，此时 $f_Z(z)=0$;
2、当 $z<0, z>a/c\ and\ z<b/d$, 相应的取值范围是 $\frac{b}{z}<v<\frac{a}{z}, c<v<d$, 则积分区域是 $[\max(b/z,c),\min(a/z,d)]$, 此时 $f_Z(z)=\int_{p(z)}^{q(z)} f(zv,v)|v|dv=\frac{q^2(z)-p^2(z)}{2(b-a)(d-c)}$ 这里 $p(z)=\max(b/z,c), q(z)=\min(a/z,d)$;
3、当 $z>0, z<a/b\ or\ z>b/c$, 此时事件是不可能事件，此时 $f_Z(z)=0$;
4、当 $z>0, z>a/b\ and\ z<b/c$, 相应的取值范围是 $\frac{a}{z}<v<\frac{b}{z}, c<v<d$, 则积分区域是 $[\max(a/z,c),\min(b/z,d)]$, 此时 $f_Z(z)=\int_{m(z)}^{n(z)} f(zv,v)|v|dv=\frac{n^2(z)-m^2(z)}{2(b-a)(d-c)}$ 这里 $m(z)=\max(a/z,c), n(z)=\min(b/z,d)$;

$$f(z)=\begin{cases}\frac{q^2(z)-p^2(z)}{2(b-a)(d-c)}, & z<0, z>a/c\ and\ z<b/d;\\ \frac{n^2(z)-m^2(z)}{2(b-a)(d-c)}, & z>0, z>a/b\ and\ z<b/c;\\ 0, & otherwise.\end{cases}$$

例 3.21 假设 Y_1, Y_2 服从二元正态分布具有参数 $\mu_{Y_1}=\mu_{Y_2}=0, \sigma^2_{Y_1}=1, \sigma^2_{Y_2}=2$, 且 $\rho=1/\sqrt{2}$. 找出线性变换 $x_1=a_{11}y_1+a_{12}y_2, x_2=a_{21}y_1+a_{22}y_2$. 使得 X_1 和 X_2 是独立的标准正态随机变量。

解答：

由题意，可知

$$\begin{cases}Y_1=\frac{a_{22}X_1-a_{12}X_2}{a_{11}a_{22}-a_{12}a_{21}},\\ Y_2=\frac{a_{11}X_2-a_{21}X_1}{a_{11}a_{22}-a_{12}a_{21}}.\end{cases}$$

则利用密度函数的变换可得

$$f_{X_1,X_2}(x_1,x_2)=f_{X_1,X_2}(\frac{a_{22}X_1-a_{12}X_2}{a_{11}a_{22}-a12a_{21}},\frac{a_{11}X_2-a_{21}X_1}{a_{11}a_{22}-a12a_{21}})|J|$$

如果 X_1 和 X_2 是独立的标准正态随机变量，那么 $f_{X_1,X_2}(x_1,x_2)=1/2\pi\exp\{-(x_1^2/2+x_2^2/2)\}$.

令 $c=\frac{1}{a_{11}a_{22}-a12a_{21}}$, 则

$$c^2[(a_{22}X_1-a_{12}X_2)^2-(a_{22}X_1-a_{12}X_2)(a_{11}X_2-a_{21}X_1)+\frac{(a_{11}X_2-a_{21}X_1)^2}{2}=x_1^2/2+x_2^2/2.$$

对应起来，可得

$$\begin{cases} c^2[a_{22}^2+a_{22}a_{21}+a_{21}^2/2]=1; \\ c^2[a_{12}^2+a_{11}a_{12}+a_{11}^2/2]=1; \\ c^2[2a_{22}a_{12}+a_{22}a_{11}+a_{12}a_{21}+2a_{11}a_{12}]=0. \end{cases}$$

显然这些方程不能够完全确定这些参数，可以先选取 (a_{11},a_{12}), 然后根据关系式确定 (a_{21},a_{22}). 比如 $\pm(1,0)$ 和 $\pm(1,-1)$; $(2/\sqrt{34},3/\sqrt{34})$ 和 $(3/\sqrt{5},-2/\sqrt{5})$.

例 3.22 假设 $X_1,\cdots,X_n$ 是独立的连续型随机变量，每个变量具有累积分布函数 F. 证明：$X_{(1)},X_{(n)}$ 的联合分布函数是

$$F(x,y)=F^n(y)-[F(y)-F(x)]^n,\ x\le y.$$

解答：

证法一、$X_{(1)},X_{(n)}$ 的联合密度函数是

$$f(x,y)=n(n-1)f(x)f(y)[F(y)-F(x)]^{n-2}.$$

事件 $\{X_{(1)}\le x, X_{(n)}\le y\}$ 对应的点集为 $D:=\{(u,v):u\le x,v\le y,u\le v\}$. 因此

$$\begin{aligned} F(x,y) &= \mathbb{P}(X_{(1)}\le x,X_{(n)}\le y) \\ &= \iint_D f(u,v)dudv \\ &= \int_{-\infty}^x\int_u^y n(n-1)f(u)f(v)[F(v)-F(u)]^{n-2}dvdu \\ &= \int_{-\infty}^x nf(u)[F(y)-F(u)]^{n-1}dvdu \\ &= -[F(y)-F(u)]|_{-\infty}^x = F^n(y)-[F(y)-F(x)]^n,\ x\le y. \end{aligned}$$

证法二、由定义，$X_{(1)},X_{(n)}$ 的联合分布函数是

$$F(x,y)=\mathbb{P}(X_{(1)}\le x,X_{(n)}\le y),$$

因为事件 $\{X_{(1)}\le x\}$ 的概率不好求，一般考虑其补事件，因为事件

$$\{X_{(1)}\le x,X_{(n)}\le y\}:=\{X_{(n)}\le y\}-\{X_{(1)}>x,X_{(n)}\le y\}.$$

因为

$$\mathbb{P}(X_{(n)}\le y)=F^n(y)$$

和

$$\mathbb{P}(X_{(1)} > x, X_{(n)} \le y) = [F(y) - F(x)]^n.$$

因此

$$F(x, y) = F^n(y) - [F(y) - F(x)]^n,\ x \le y.$$

例 3.23 设二维随机变量 (X, Y) 的联合密度函数为:

$$f(x,y) = \begin{cases} 3x, & 0 < x < 1, 0 < y < x; \\ 0, & otherwise. \end{cases}$$

求 $Z = X - Y$ 的密度函数.

解答:

解法一、直接计算分布函数，然后求导。显然事件 $\{X - Y \le z\}$ 的概率是密度函数在 $x = y$ 和 $x - y = z$ 夹成的区域之间的积分。

$$\begin{aligned} F_Z(z) &= \mathbb{P}(X - Y \le z) \\ &= \int_0^z \int_0^x 3x dy dx + \int_z^1 \int_{x-z}^x 3x dy dx \\ &= 3/2z - 1/2z^3 \end{aligned}$$

因此

$$f_Z(z) = F_Z'(z) = \begin{cases} 3/2(1 - z^2), & 0 < x < 1, 0 < y < x; \\ 0, & otherwise. \end{cases}$$

解法二、利用变换，令 $V = Y$ 则 $\begin{cases} X = Z + V; \\ Y = V. \end{cases}$ 原来的支集 $\{(x, y) : 0 < x < 1, 0 < y < x\}$, 则变换之后的区域是

$$\{(z, v) : 0 < z + v < 1, 0 < v < z + v\} = \{(z, v) : 0 < z + v < 1, 0 < z\}.$$

因此

$$f_{Z,V}(z, v) = f(z + v, v)|J| = \begin{cases} 3(z + v), & 0 < z + v < 1, 0 < z; \\ 0, & otherwise. \end{cases}$$

则 Z 的边际密度为

$$f_Z(z) = \int_0^{1-z} 3(z + v) dv = \begin{cases} 3/2(1 - z^2), & 0 < z < 1; \\ 0, & otherwise. \end{cases}$$

例 3.24 设二维随机变量 (X, Y) 的联合密度函数为:

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

1. 问 X 与 Y 是否相互独立?

2. 求 $Z = X + Y$ 的密度函数.

解答:

1、由定义 X 的边际密度

$$f_X(x) = \int_0^{\infty} \frac{1}{2}(x+y)e^{-(x+y)}dy = \frac{x+1}{2}e^{-x}, x > 0.$$

同理，Y 的边际 $f_Y(y) = \frac{y+1}{2}e^{-y}, y > 0$. 因为 $f(x,y) \neq f_X(x)f_Y(y)$, 所以不独立。2、解法一、先求分布函数然后求导，由题意，

$$\begin{aligned} F_Z(z) &= \mathbb{P}(X+Y \leq z) \\ &= \int_0^z \int_0^{z-x} \frac{1}{2}(x+y)e^{-(x+y)}dydx \\ &= \int_0^z \int_x^z \frac{1}{2}ue^{-u}dudx \\ &= \frac{1}{2} - \frac{2+2z+z^{2-z}}{e} \end{aligned}$$

因此 Z 的密度函数

$$f_Z(z) = F_Z'(z) = \begin{cases} \frac{z^2}{2}e^{-z}, & 0 < v < z; \\ 0, & otherwise. \end{cases}$$

解法二、利用换元，假设 $V = Y$, 则 $\begin{cases} X = Z - V; \\ Y = V. \end{cases}$ 原来的支集 $\{(x,y) : 0 < x, 0 < y\}$, 则变换之后的区域是

$$\{(z,v) : 0 < z - v, 0 < v\} = \{(z,v) : 0 < v < z\}.$$

因此 (Z,V) 的联合密度函数

$$f_{Z,V}(z,v) = f(z-v,v)|J| = \begin{cases} 1/2e^{-z}, & 0 < v < z; \\ 0, & otherwise. \end{cases}$$

则 Z 的边际密度为

$$f_Z(z) = \int_0^z 1/2ze^{-z}dv = \begin{cases} \frac{z^2}{2}e^{-z}, & 0 < z; \\ 0, & otherwise. \end{cases}$$

例 3.25 设 X 与 Y 是相互独立的随机变量，且在 $[0,1]$ 上服从均匀分布，分别求随机变量 $\xi = \max(X,Y), \eta = \min(X,Y)$ 的概率密度.

答案:

$\xi = \max(X,Y)$ 的概率密度为

$$f_\xi(x) = \begin{cases} 2x, & 0 < x < 1; \\ 0, & otherwise. \end{cases}$$

$\eta=\min(X,Y)$ 的概率密度为

$$f_\eta(x)=\begin{cases}2(1-x), & 0<x<1;\\ 0, & otherwise.\end{cases}$$

例 3.26 设 $X_1,\cdots,X_n$ 是相互独立的 n 个独立同分布的随机变量，X_i 都服从参数为 β 的指数分布。若 $Y=\min(X_1,\cdots,X_n)$, 求 Y 的分布.

答案:

$Y=\min(X_1,\cdots,X_n)$, 求 Y 的分布函数为

$$F_Y(y)=1-\mathbb{P}(Y>y)=\begin{cases}1-e^{-n\beta y}, & 0<y;\\ 0, & otherwise.\end{cases}$$

Y 的概率密度为

$$f_Y(y)=\begin{cases}n\beta e^{n\beta y}, & 0<y;\\ 0, & otherwise.\end{cases}$$

例 3.27 假设 A,B,C 是服从 $(0,1)$ 上均匀分布的随机变量，试求方程

$$Ax^2+Bx+C=0$$

有实根的概率。

解答:

定义事件 $A:=\{方程Ax^2+Bx+C=0有实根\}$, 则可知 $A:=\{B^2>4AC\}$. 因此

$$\begin{aligned}
\mathbb{P}(A) &= \mathbb{P}(B^2\geq 4AC)\\
&= \iiint\limits_{b^2\geq 4ac} 1dadbdc\\
&= \int_0^1\int_0^1\int_{2\sqrt{ac}}^1 1dbdadc\\
&= 1-2\iint\limits_{4ac\leq 1}\sqrt{ac}dadc\\
&= 1-2[\int_0^1\int_0^{1/4}\sqrt{ac}dadc+\int_{1/4}^1\int_0^{1/4a}\sqrt{ac}dcda]\\
&= \frac{8-3\ln 2}{9}
\end{aligned}$$

quiz 答案:

quiz 1

1、≈ 0.5; 2、0.1； 3、19/36.

quiz 2

1、1/37, 36/37; 2、0.98； 3、0.085. quiz 3

1、$P(A\cup B)=0.9, P(A|\bar{B})=0.75; P(A\cup B)=0.72, P(A|\bar{B})=0.3$；2、$A=1$; 3、略；4、$\binom{k+r-1}{r}p^r(1-p)^k$; 5、$(1-p)^k p$; 6、19/27.

quiz 4

1、$1-\sum\limits_{k=0}^{20}\frac{10^k}{k!}e^{-10}, \sum\limits_{k=0}^{15}\frac{10^k}{k!}e^{-10}$；2、2；3、略；4、略；5、19/27；6、2/3；7、$p_0=0.01, p_1=0.18, p_2=0.82$; 8、略，9、

$$f_X(x)=\begin{cases}\frac{1}{a}, & 0<x<a;\\ 0, & otherwise.\end{cases}$$

和

$$F_X(x)=\begin{cases}0, & x<0;\\ \frac{x}{a}, & 0<x<a;\\ 1, & x>a.\end{cases}$$

quiz 5

1、略；2、3、略；0.59；24/65；4、略；5、-1/2；略；略；6、

$$f_Y(y)=\begin{cases}\frac{1}{2\sqrt{\pi(y-1}}e^{-(y-1)/4}, & 1<y;\\ 0, & otherwise.\end{cases}$$

7、略；8、当 $c>0$ 时

$$f_Y(y)=\begin{cases}\frac{1}{c(b-a)}, & a<(y-d)/c<b;\\ 0, & otherwise.\end{cases}$$

当 $c<0$ 时

$$f_Y(y)=\begin{cases}-\frac{1}{c(b-a)}, & a<(y-d)/c<b;\\ 0, & otherwise.\end{cases}$$

quiz 6

1、略；2、

quiz 7

1、略；2、$k=4$;

$$f_X(x)=\begin{cases}2xe^{-x^2}, & 0<x;\\ 0, & otherwise.\end{cases}$$

和

$$f_Y(y)=\begin{cases}2ye^{-y^2}, & 0<y;\\ 0, & otherwise.\end{cases}$$

独立；

$$f_{X|Y}(x|y)=\begin{cases}2xe^{-x^2}, & 0<x;\\ 0, & otherwise.\end{cases}$$

3、$k=3$;

$$f_{X|Y}(x|y)=\begin{cases}\frac{2x}{1-y^2}, & y<x<1;\\ 0, & otherwise.\end{cases}$$

0; 4、1/3. quiz 8

1、

$$f_Z(z)=\begin{cases}3/2(1-z^2), & 0<z<1;\\ 0, & otherwise.\end{cases}$$

2、不独立

$$f_Z(z)=\begin{cases}\frac{z^2}{2}e^{-z}, & 0<z;\\ 0, & otherwise.\end{cases}$$

3、分布函数

$$F_Y(y)=1-\mathbb{P}(Y>y)=\begin{cases}1-e^{-n\beta y}, & 0<y;\\ 0, & otherwise.\end{cases}$$

Y 的概率密度为

$$f_Y(y)=\begin{cases}n\beta e^{n\beta y}, & 0<y;\\ 0, & otherwise.\end{cases}$$

4、$\xi=\max(X,Y)$ 的概率密度为

$$f_\xi(x)=\begin{cases}2x, & 0<x<1;\\ 0, & otherwise.\end{cases}$$

$\eta=\min(X,Y)$ 的概率密度为

$$f_\eta(x)=\begin{cases}2(1-x), & 0<x<1;\\ 0, & otherwise.\end{cases}$$

第 4 章
第四章：随机变量的特征数

4.1 期望、方差、协方差、条件期望

1. 数学期望（离散）：如果级数 $\sum_{k=1}^{\infty}|x_k|p_k<\infty$, 则称

$$\mathbb{E}(X)=\sum_{k=1}^{\infty}x_kp_k$$

为 r.v. X 的期望；

数学期望（连续型）：假设密度函数为 $f(x)$ 且 $\int_{-\infty}^{\infty}|x|f(x)dx<\infty$, 则称

$$\mathbb{E}(X)=\int_{-\infty}^{\infty}xf(x)dx$$

为 r.v. X 的期望.

2. 常见随机变量的期望：Poisson 分布：$\mathbb{E}(X)=\lambda$; 二项分布：$\mathbb{E}(X)=np$; 几何分布：$\mathbb{E}(X)=\frac{1}{p}$; 负二项分布：$\mathbb{E}(X)=\frac{r}{p}$; 超几何分布：$\mathbb{E}(X)=\frac{nm}{N}$; 均匀分布：$\mathbb{E}(X)=\frac{a+b}{2}$; 正态分布：$\mathbb{E}(X)=\mu$; 指数分布：$\mathbb{E}(X)=\frac{1}{\lambda}$; Γ 分布：$\mathbb{E}(X)=\frac{\alpha}{\lambda}$; 柯西分布：不存在；$\beta$ 分布：$\mathbb{E}(X)=\frac{a}{a+b}$.

3. 函数的数学期望：函数为 $Y=g(X)$ 的期望为（离散型）

$$\mathbb{E}(Y)=\mathbb{E}(g(X))=\sum_{k=1}^{\infty}g(x_k)p_k;$$

或（连续型）

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

4. 对于二元甚至 n 元随机变量的函数上述表达形式依然成立.

5. 期望的性质：

   - 如果 $a\le X\le b$, 则 $a\le\mathbb{E}(X)\le b$;
   - 对于常数 c, 则 $\mathbb{E}(cX)=c\mathbb{E}(X)$;

- 对于任意两个随机变量 X, Y, 则 $\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y)$;
- 如果随机变量 X, Y 相互独立，则 $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$;
- 推广到一般情形。

6. **定理：Markov 不等式**　假设随机变量 X 满足 $\mathbb{P}(X \geq 0) = 1$, 且 $\mathbb{E}(X)$ 存在，则
$$\mathbb{P}(X > t) \leq \frac{\mathbb{E}(X)}{t}.$$

7. 对 r.v. X，如果
$$Var(X) = D(X) = \mathbb{E}[X - \mathbb{E}(X)]^2$$
存在，则称 $D(X)$ 为 X 的方差，$\sqrt{D(X)}$ 为 X 的标准差。

8. 常见随机变量的方差：Poisson 分布: $D(X) = \lambda$; 二项分布: $D(X) = np(1-p)$; 几何分布：$D(X) = \frac{1-p}{p^2}$; 负二项分布：$D(X) = \frac{r(1-p)}{p^2}$; 超几何分布：$D(X) = np(1-p)(1-\frac{n-1}{N-1})$, 其中 $p = \frac{m}{N}$; 均匀分布：$D(X) = \frac{(b-a)^2}{12}$; 指数分布 $D(X) = \frac{1}{\lambda^2}$; 正态分布：$D(X) = \sigma^2$; Γ 分布：$D(X) = \frac{\alpha}{\lambda^2}$; 柯西分布：不存在；$\beta$ 分布：$D(X) = \frac{ab}{(a+b)^2(a+b+1)}$.

9. 方差的性质
   - $X = c$ 当且仅当 $D(X) = 0$;
   - 假设 c 为常数，则 $D(cX) = c^2 D(X)$;
   - 对任意随机变量 X, Y:
   $$D(X+Y) = D(X) + D(Y) + 2\mathbb{E}[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))];$$
   - 对独立随机变量 X, Y:
   $$D(X+Y) = D(X) + D(Y);$$

10. **定理：Chebyshev 不等式**　假设随机变量 X 满足 $\mathbb{E}(X) = \mu$ 和 $D(X) = \sigma^2$ 都存在，则
$$\mathbb{P}(|X - \mu| > \varepsilon) = \frac{\sigma^2}{\varepsilon^2}.$$

11. 对 r.v. X, Y，如果相应的方差都存在，则称
$$Cov(X, Y) = \mathbb{E}[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))]$$
为 r.v. X, Y 的协方差。

12. 协方差的性质

- 对独立随机变量 X, Y 有 $Cov(X,Y) = 0$;
- $Cov(X,Y) = Cov(Y,X)$; $D(X) = Cov(X,X)$;
- $Cov(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$, 因此 $D(X+Y) = D(X) + D(Y) + 2Cov(X,Y)$.
- $Cov(aX, bY) = abCov(X,Y)$.
- $Cov(X_1 + X_2, Y) = Cov(X_1, Y) + Cov(X_2, Y)$.

13. **协方差的双线性** 假设随机变量 $U = a + \sum_{i=1}^n a_i X_i$ 和 $V = b + \sum_{j=1}^m b_j X_j$, 有

$$Cov(U,V) = \sum_{i=1}^n \sum_{j=1}^m a_i b_j Cov(X_i, Y_j).$$

14. 对 r.v. X, Y，如果相应的方差都存在，则称

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}$$

为 r.v. X, Y 的相关系数。

15. **定理** 相关系数 $|\rho| \le 1$, 如果 $|\rho| = 1$, 则有 $Y = aX + b$.

16.
- 高阶原点矩和中心距: $\mathbb{E}(X^k)$ 和 $\mathbb{E}(X - \mathbb{E}(X))^k$.
- 多维随机变量 $(X_1, \cdots, X_n)$ 的协方差矩阵

$$C = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{bmatrix}$$

其中 $c_{ij} = \mathbb{E}[(X_i - \mathbb{E}(X_i))(X_j - \mathbb{E}(X_j))]$.

- 多维正态密度:

$$f(x_1, x_2, \cdots, x_n) = \frac{1}{(2\pi)^{n/2}|C|^{1/2}} \exp\{-\frac{1}{2}(X-\mu)' C^{-1} (X-\mu)\}.$$

17. 条件期望

- 在给定 $X = x$ 的条件下，Y 的条件期望:

$$\begin{aligned} \mathbb{E}(Y|X=x) &= \sum_y y p_{Y|X}(y|x) \text{ (离散情形)} \\ \mathbb{E}(Y|X=x) &= \int_{-\infty}^{\infty} y f_{Y|X}(y|x) dy \text{ (连续情形)} \end{aligned}$$

- **定理：重期望公式**：$\mathbb{E}(Y)=\mathbb{E}[\mathbb{E}(Y|X)]$；和 $D(Y)=D(\mathbb{E}(Y|X))+\mathbb{E}(D(Y|X))$.

Quiz 9

1. 设随机变量 X 的联合密度函数为：

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

求 $\mathbb{E}(2X+5), D(X)$，（这里 $D(X)$ 是指 X 的方差，符号等同 $Var(X)$.）

解答：

由期望的性质，有

$$\begin{aligned} \mathbb{E}(2X+5) &= 2\mathbb{E}(X)+5 \\ &= 2\int_0^\infty xe^{-x}dx+5 \\ &= 7 \end{aligned}$$

同理，$\mathbb{E}(X)=1$ 和 $\mathbb{E}(X^2)=\int_0^\infty x^2e^{-x}dx=2$，因此由定义

$$D(X)=\mathbb{E}(X^2)-[\mathbb{E}(X)]^2=1.$$

2. 设随机变量 X 的联合密度函数为：

$$f(x)=\begin{cases} \alpha x(1-x), & 0<x<1; \\ 0, & otherwise. \end{cases}$$

求 $\alpha, \mathbb{E}(X), D(X)$.

解答：

由密度函数的性质，可知

$$\int_0^1 \alpha x(1-x)dx = a/6 = 1,$$

因此 $a=6$. 由期望的性质，有

$$\begin{aligned} \mathbb{E}(X) &= \int_0^1 6x^2(1-x)dx=\frac{1}{2} \\ \mathbb{E}(X^2) &= \int_0^1 6x^3(1-x)dx=\frac{3}{10} \end{aligned}$$

因此由定义

$$D(X)=\mathbb{E}(X^2)-[\mathbb{E}(X)]^2=\frac{1}{20}.$$

3. 某设备由三大部件构成，设备运转时，各部件需调整的概率为 0.1, 0.2, 0.3，若各部件的状态相互独立，求同时需调整的部件数 X 的期望与方差.

解答:
由题意随机变量 X 的分布列为

X	0	1	2	3
p	0.504	0.398	0.092	0.006

因此

$$\begin{aligned}\mathbb{E}(X) &= 0\times 0.504+1\times 0.398+2\times 0.092+3\times 0.006=0.6\\ \mathbb{E}(X^2) &= 0\times 0.504+1\times 0.398+4\times 0.092+9\times 0.006=0.82.\end{aligned}$$

由定义

$$D(X)=\mathbb{E}(X^2)-[\mathbb{E}(X)]^2=0.46.$$

4. 设离散型二维随机变量 (X,Y) 在点 $(1,1),(\frac{1}{2},\frac{1}{4}),(-\frac{1}{2},-\frac{1}{4}),(-1,-1)$ 取值的概率均为 1/4，求 $\mathbb{E}(X),\mathbb{E}(Y),D(X),D(Y),\mathbb{E}(XY)$.

解答:
由定义

$$\begin{aligned}\mathbb{E}(X) &= 1\times 1/4+\frac{1}{2}\times 1/4-\frac{1}{2}\times 1/4-1\times 1/4=0\\ \mathbb{E}(X^2) &= 1\times 1/4+\frac{1}{4}\times 1/4+\frac{1}{4}\times 1/4+1\times 1/4=\frac{5}{8}\end{aligned}$$

因此 $D(X)=\frac{5}{8}$. 同理 $\mathbb{E}(Y)=0,D(Y)=\frac{17}{32}$. 由期望的性质

$$\mathbb{E}(XY)=1\times 1/4+\frac{1}{8}\times 1/4+\frac{1}{8}\times 1/4+1\times 1/4=\frac{9}{16}.$$

例 4.1 假设 X,Y 均服从正态分布则

1. $X+Y$ 一定服从正态分布；

2. X 和 Y 不相关与独立等价;

3. (X,Y) 一定服从正态分布；

4. $(X,-Y)$ 未必服从正态分布；

解答：

第四个答案是正确的。分析：假如 X 服从正态分布，显然 $Y=-X$ 也服从正态分布，但 $X+Y=0$ 显然不服从正态分布；对于第二个论述，假设 X 为标准正态随机变量，S 为独立于 X 且分别以概率 $1/2$ 取值 1 和 -1，利用作业第 60 题结论，可知 X,Y 不相关但也不独立。P.s. 如果 (X,Y) 服从二元正态，那么不相关和独立是等价的；第三问 (X,Y) 显然不一定服从二元正态分布。

例 4.2： 在 n 次独立重复实验中，假设 X,Y 分别表示成功的次数和失败的次数，则 X,Y 的相关系数为：

1. -1；
2. 0;
3. 1/2；
4. 1；

解答：

-1, 因为 $Y=n-X$.

例 4.3 设 (X,Y) 的联合密度函数为

$$f(x,y)=\begin{cases}2\frac{e^{-2x}}{x}, & 0<y<x<\infty;\\ 0, & otherwise.\end{cases}$$

试求 $Cov(X,Y)$.

解答：

因为 $Cov(X,Y)=\mathbb{E}(XY)-\mathbb{E}(X)\mathbb{E}(Y)$, 只需求 $\mathbb{E}(XY),\mathbb{E}(X),\mathbb{E}(Y)$ 即可。由定义

$$
\begin{aligned}
\mathbb{E}(XY) &= \iint_{\mathbb{R}^2} xyf(x,y)dxdy \\
&= \int_0^{\infty}\int_0^{x} 2xy\frac{e^{-2x}}{x}dydx \\
&= \int_0^{\infty} x^2e^{-2x}dx \\
&= \frac{1}{4}
\end{aligned}
$$

同理

$$
\begin{aligned}
\mathbb{E}(X) &= \iint_{\mathbb{R}^2} xf(x,y)dxdy \\
&= \int_0^{\infty}\int_0^{x} 2x\frac{e^{-2x}}{x}dydx \\
&= \int_0^{\infty} 2e^{-2x}dx \\
&= \frac{1}{2}
\end{aligned}
$$

和

$$
\begin{aligned}
\mathbb{E}(Y) &= \iint_{\mathbb{R}^2} yf(x,y)dxdy \\
&= \int_0^{\infty}\int_0^{x} 2y\frac{e^{-2x}}{x}dydx \\
&= \int_0^{\infty} xe^{-2x}dx \\
&= \frac{1}{4}
\end{aligned}
$$

因此 $Cov(X,Y) = \frac{1}{8}$.

例 4.4 设 (X,Y) 的联合密度函数为

$$
f(x,y) = \begin{cases} \frac{e^{-(y+x/y)}}{y}, & 0 < y < \infty, 0 < x < \infty; \\ 0, & otherwise. \end{cases}
$$

试求 $Cov(X,Y)$ 和 $\mathbb{E}(X^2|Y=y)$.

解答：

因为 $Cov(X,Y)=\mathbb{E}(XY)-\mathbb{E}(X)\mathbb{E}(Y)$, 只需求 $\mathbb{E}(XY),\mathbb{E}(X),\mathbb{E}(Y)$ 即可。由定义

$$\begin{aligned}\mathbb{E}(XY) &= \iint_{\mathbb{R}^2} xyf(x,y)dxdy \\ &= \int_0^\infty\int_0^\infty xy\frac{e^{-(y+x/y)}}{y}dxdy \\ &= \int_0^\infty y^2e^{-y}dy \\ &= 2\end{aligned}$$

同理

$$\begin{aligned}\mathbb{E}(X) &= \iint_{\mathbb{R}^2} xf(x,y)dxdy \\ &= \int_0^\infty\int_0^\infty x\frac{e^{-(y+x/y)}}{y}dxdy \\ &= \int_0^\infty ye^{-y}dy \\ &= 1\end{aligned}$$

和

$$\begin{aligned}\mathbb{E}(Y) &= \iint_{\mathbb{R}^2} yf(x,y)dxdy \\ &= \int_0^\infty\int_0^\infty y\frac{e^{-(y+x/y)}}{y}dxdy \\ &= \int_0^\infty ye^{-y}dy \\ &= 1\end{aligned}$$

因此 $Cov(X,Y)=1$.

由定义 $\mathbb{E}(X^2|Y=y)=\int_{-\infty}^{\infty}x^2f_{X|Y}(x|y)dx$, 因此我们需要条件密度 $f_{X|Y}(x|y)$. 随机变量 Y 的边际密度函数为

$$\begin{aligned}f_Y(y) &= \int_{-\infty}^{\infty}f(x,y)dx \\ &= \begin{cases}\int_0^\infty \frac{e^{-(y+x/y)}}{y}dx, & 0<y; \\ 0, & otherwise.\end{cases} \\ &= \begin{cases}e^{-y}, & 0<y; \\ 0, & otherwise.\end{cases}\end{aligned}$$

因此，在给定条件 $Y=y$ 时，X 的条件密度为

$$f_{X|Y}(x|y)=\begin{cases}\frac{e^{-x/y}}{y}, & 0<x; \\ 0, & otherwise.\end{cases}$$

于是

$$\begin{aligned}\mathbb{E}(X|Y=y) &= \int_{-\infty}^{\infty} x^2 f_{X|Y}(x|y)dx \\ &= \int_0^{\infty} x^2 \frac{e^{-x/y}}{y} dxdy \\ &= 2y^2\end{aligned}$$

例 4.5 设 (X,Y) 服从二元正态 $N(0,0,1,1,\rho)$ 试求 $X-Y$ 和 XY 之间的协方差以及相关系数。

解答：

由协方差的性质：

$$Cov(X-Y,XY)=Cov(X,XY)-Cov(Y,XY)=\mathbb{E}(X^2Y)-\mathbb{E}(XY^2)=0.\ (why?)$$

练习 4.6 假设 U 服从 $(-2,2)$ 上的均匀分布，定义 X 和 Y 如下

$$X=\begin{cases}-1, & U<-1;\\ 1, & U>-1.\end{cases} \quad Y=\begin{cases}-1, & U<1;\\ 1, & U>1.\end{cases}$$

试求 $Var(X+Y)$.

答案： 2

练习：4.7 设 (X,Y) 的联合密度函数为

$$f(x,y)=\begin{cases}\frac{x(1+3y^2)}{4}, & 0<y<1, 0<x<2;\\ 0, & otherwise.\end{cases}$$

试求 $\mathbb{E}(\frac{Y}{X})$.

答案： $\frac{5}{8}$.

例 4.8 在一个有 n 个人参加的舞会，每个人带了一份礼物，且假定每个人带的礼物各不相同。晚会期间人们把礼物放在一起，且每人随机的抽取一件。试求抽中自己礼品的人数 X 的期望和方差。

解答：

记随机变量

$$X_i = \begin{cases} 1, & \text{第 i 个人恰好拿到自己的礼物;} \\ 0, & \text{第 i 个人没有拿到自己的礼物.} \end{cases}$$

则可知 $X_i, i = 1, 2, \cdots, n$ 是同分布的，但不独立。其共同分布为

$$\mathbb{P}(X_i = 1) = \frac{1}{n}, \mathbb{P}(X_i = 0) = \frac{n-1}{n}, \quad i = 1, 2, \cdots, n.$$

由此，可知 $\mathbb{E}(X_i) = \frac{1}{n}$ 和 $Var(X_i) = \frac{n-1}{n^2}$. 此外 $X = X_1 + X_2 + \cdots + X_n$, 可知

$$\mathbb{E}(X) = \mathbb{E}(X_1) + mathbbE(X_2) + \cdots + \mathbb{E}(X_n) = 1.$$

由随机变量和的方差公式可知

$$Var(X) = \sum_{i=1}^{n} Var(X_i) \sum_{i=1}^{n} \sum_{j=i+1}^{n} 2Cov(X_i, X_j).$$

而

$$Cov(X_i, X_j) = \mathbb{E}(X_i X_j) - \mathbb{E}(X_i)\mathbb{E}(X_j) = \frac{1}{n(n-1)} - \frac{1}{n^2} = \frac{1}{n^2(n-1)}.$$

将其带入前面等式，可知

$$Var(X) = \frac{n-1}{n} - 2\binom{n}{2}\frac{1}{n^2(n-1)} = 1.$$

例 4.9 （**Quiz9 B-Q2**） 设二维随机变量 (X, Y) 服从以点 $(0,1), (1,0), (1,1)$ 为顶点的三角形区域上的均匀分布。试求 $\mathbb{E}(X+Y)$ 和 $D(X+Y)$.

解答：

记所围成的三角形区域为 D, 因为 D 的面积为 $1/2$, 因此 (X,Y) 的联合密度函数为

$$f(x,y) = \begin{cases} 2, & (x,y) \in D; \\ 0, & otherwise. \end{cases}$$

(可以直接求，但请注意积分上下限。这里先求边际，再求期望方差)

当 $0 < x < 1$, 有

$$f_X(x) = \int_{1-x}^{1} 2dy = 2x.$$

同理，当 $0<y<1$ 时，有

$$f_Y(y)=\int_{1-y}^{1}2dx=2y.$$

利用期望和方差的定义可知

$$\mathbb{E}(X)=\mathbb{E}(Y)=\frac{2}{3},\ \ Var(X)=Var(Y)=\frac{1}{18}.$$

因为

$$\mathbb{E}(XY)=\int_0^1\int_{1-x}^{1}2xydydx=\frac{5}{12}.$$

由此得

$$Cov(X,Y)=\mathbb{E}(XY)-\mathbb{E}(X)\mathbb{E}(Y)=-\frac{1}{36}.$$

最终

$$\begin{aligned}\mathbb{E}(X+Y)&=\mathbb{E}(X)+\mathbb{E}(Y)=\frac{4}{3}\\ D(X+Y)&=D(X)+D(Y)+2Cov(X,Y)=\frac{1}{18}.\end{aligned}$$

例 4.10 （**Quiz9 Q1**） 设二维随机变量 (X,Y) 的联合密度函数为：

$$f(x,y)=\begin{cases}3x, & 0<y<x<1;\\ 0, & otherwise.\end{cases}$$

求 X 与 Y 的期望、方差、协方差和相关系数.

解答：

由定义，需要先计算 X,Y,X^2,Y^2,XY 的期望

$$\begin{aligned}\mathbb{E}(X)&=\int_0^1\int_0^x 3x^2dydx=\frac{3}{4};\\ \mathbb{E}(X^2)&=\int_0^1\int_0^x 3x^3dydx=\frac{3}{5};\\ \mathbb{E}(Y)&=\int_0^1\int_0^x 3xydydx=\frac{3}{8};\\ \mathbb{E}(Y^2)&=\int_0^1\int_0^x 3xy^2dydx=\frac{1}{5};\\ \mathbb{E}(XY)&=\int_0^1\int_0^x 3x^2ydydx=\frac{3}{10}.\end{aligned}$$

则可得

$$\begin{aligned} Var(X) &= \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = \frac{3}{80}; \\ Var(Y) &= \mathbb{E}(Y^2) - (\mathbb{E}(Y))^2 = \frac{19}{320}; \\ Cov(X,Y) &= \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = \frac{3}{160}. \end{aligned}$$

最后可得 X 和 Y 的相关系数为

$$Corr(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}} = \frac{3}{\sqrt{57}} \approx 0.397.$$

练习 4.11 设二维随机变量 (X,Y) 的联合密度函数为：

$$f(x,y) = \begin{cases} 1, & |y| < x, 0 < x < 1; \\ 0, & otherwise. \end{cases}$$

求 X 与 Y 的期望、协方差.

答案： $\mathbb{E}(X) = 2/3, \mathbb{E}(Y) = 0, Cov(X,Y) = 0.$

练习 4,12 邮局里有 A、B、C 三名顾客，假定邮局对每位顾客的服务时间服从参数为 λ 的指数分布，对 A 和 B 立即开始服务，在对 A 或 B 服务结束后立即对 C 服务，对 A、B 服务所需要的时间相互独立，求（1）C 在邮局等待的时间的期望；（2）C 在邮局逗留时间的期望.

答案： $\frac{1}{2\lambda}, \frac{3}{2\lambda}.$

例 4.13 设 a 是区间 $(0,1)$ 上的一个定点，随机变量 X 服从 $(0,1)$ 上的均匀分布，用 Y 表示 X 到 a 点的距离，问 a 取何值时 X 和 Y 不相关。

解答：
由题意可知 $X \sim U(0,1)$, 因此 $\mathbb{E}(X) = \frac{1}{2}$. 又因为

$$\begin{aligned} \mathbb{E}(Y) &= \int_0^a (a-x)dx + \int_a^1 (x-a)dx = a^2 - a + \frac{1}{2}; \\ \mathbb{E}(XY) &= \int_0^a x(a-x)dx + \int_a^1 x(x-a)dx = \frac{a^3}{3} - \frac{a}{2} + \frac{1}{3}. \end{aligned}$$

可得

$$Cov(X,Y)=\frac{a^3}{3}-\frac{a^2}{2}+\frac{1}{12},$$

协方差 $Cov(X,Y)=0$ 等价于

$$4a^3-6a^2+1=0.$$

因式分解得到

$$(2a-1)(2a^2-2a-1)=0$$

因此当且仅当 $a=0.5$ 时 X 和 Y 不相关。

练习 4.14 设二维随机变量 (X,Y) 的联合密度函数如下：

1.
$$f_1(x,y)=\begin{cases}6xy^2, & 0<x<1,0<y<1;\\ 0, & otherwise.\end{cases}$$

2.
$$f_2(x,y)=\begin{cases}\frac{x+y}{8}, & 0<x<2,0<y<2;\\ 0, & otherwise.\end{cases}$$

求 (X,Y) 的协方差矩阵.

练习 4.15 设二维随机变量 (X,Y) 的联合密度函数为：

$$f(x,y)=\begin{cases}\frac{6}{7}(x^2+xy/2), & 0<x<1,0<y<2;\\ 0, & otherwise.\end{cases}$$

求 X 与 Y 的协方差和相关系数.

答案： $Cov(X,Y)=-\frac{1}{147}, Corr(X,Y)=-\frac{\sqrt{15}}{69}\approx -0.05613.$

例 4.16 如果 X 和 Y 同分布但不一定独立，且方差存在。证明

$$Cov(X+Y,X-Y)=0.$$

解答：

因为

$$Cov(X+Y, X-Y) = \mathbb{E}(X^2) - \mathbb{E}(Y^2) = 0.$$

例 4.17 设随机变量 $X_1, X_2, \cdots, X_n$ 相互独立，且都服从 $(0, \theta)$ 上的均匀分布。记

$$Y = \max\{X_1, X_2, \cdots, X_n\},\ Z = \max\{X_1, X_2, \cdots, X_n\}.$$

试求 $\mathbb{E}(Y)$ 和 $\mathbb{E}(Z)$.

解答：

记 X_1 的密度函数和分布函数分别为

$$f_1(x) = \begin{cases} \frac{1}{\theta}, & 0 < x < \theta; \\ 0, & otherwise. \end{cases} \quad F_1(x) = \begin{cases} 0, & x < 0; \\ \frac{x}{\theta}, & 0 < x < \theta \\ 1, & x > \theta. \end{cases}$$

则当 $0 < x < \theta$ 时，Y 和 Z 的密度函数分别为

$$\begin{aligned} f_Y(x) &= nf_1(x)[F_1(x)]^{n-1} = \frac{n}{\theta}(\frac{x}{\theta})^{n-1}; \\ f_Z(x) &= nf_1(x)[1 - F_1(x)]^{n-1} = \frac{n}{\theta}(1 - \frac{x}{\theta})^{n-1}. \end{aligned}$$

所以

$$\begin{aligned} \mathbb{E}(Y) &= \frac{n}{\theta^n}\int_0^\theta x^n dx = \frac{n\theta}{n+1}; \\ \mathbb{E}(Z) &= \frac{n}{\theta^n}\int_0^\theta x(\theta - x)^{n-1} dx = \frac{\theta}{n+1}. \end{aligned}$$

例 4.18，作业 60 假设随机变量 Y 的密度关于 x 轴对称，令 $X = SY$, 其中 S 是另一个独立随机变量，以概率 $\frac{1}{2}$ 分别取值 1 和-1. 证明 $Cov(X, Y) = 0$, 但 X 和 Y 不独立。

解答：

由定义 $Cov(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$. 假设随机变量 Y 的密度函数和分布函数分别为 $f_Y(y)$ 和 $F_Y(y)$, 因为 Y 的密度关于 x 轴对称, 可知 $yf_Y(y)$ 是一个奇函数。可得

$$\mathbb{E}(Y) = \int_{\mathbb{R}} yf_Y(y)dy = 0,$$

另外，利用 S 和 Y 的独立性，可知 $\mathbb{E}(X)=\mathbb{E}(S)\mathbb{E}(Y)=0$ 和 $\mathbb{E}(XY)=\mathbb{E}(S)\mathbb{E}(Y^2)=0$，由此可得 $Cov(X,Y)=0$.

下证 X 和 Y 不独立。注意证明两个随机变量独立需要证明联合分布满足 $F(x,y)=F_X(x)F_Y(y)$ 或者联合密度满足 $f(x,y)=f_X(x)f_Y(y)$ 即可，一个证明独立的简单有效的方法是证明联合分布或联合密度可以写成分离形式。反正，证明不独立只需找出一个点，上述等式不成立即可。为了避免讨论 x,y 的大小，我们考虑 $x=y$，只需找出 $x>0$ 使得 $F(x,x)\neq F_X(x)F_Y(x)$ 即可证明 X 和 Y 不独立。由定义对 $x>0$ 有

$$
\begin{aligned}
F_X(x) &= \mathbb{P}(X\le x)\\
&= \mathbb{P}(X\le x|S=1)\mathbb{P}(S=1)+\mathbb{P}(X\le x|S=-1)\mathbb{P}(S=-1)(\text{全概率公式})\\
&= \frac{1}{2}(\mathbb{P}(Y\le x)+\mathbb{P}(Y\ge -x))(\text{对称性})\\
&= F_Y(x)
\end{aligned}
$$

类似的，联合分布函数满足

$$
\begin{aligned}
F(x,x) &= \mathbb{P}(X\le x, Y\le x)\\
&= \mathbb{P}(X\le x, Y\le x|S=1)\mathbb{P}(S=1)+\mathbb{P}(X\le x, Y\le x|S=-1)\mathbb{P}(S=-1)(\text{全概率公式})\\
&= \frac{1}{2}(\mathbb{P}(Y\le x)+\mathbb{P}(-x\le Y\le x))(\text{对称性})\\
&= \frac{3}{2}F_Y(x)-\frac{1}{2}
\end{aligned}
$$

因此我们只需找出点 a 使得 $\frac{3}{2}F_Y(a)-\frac{1}{2}\neq(F_Y(a))^2$ 即可。由于一元二次方程 $\frac{3}{2}x-\frac{1}{2}=x^2$ 至多有两个解，而对连续型随机变量 $F_Y(x)$ 可以取遍 $(1/2,1)$ 上的任意值。所以肯定存在 a 使得 $\frac{3}{2}F_Y(a)-\frac{1}{2}\neq(F_Y(a))^2$.

练习 4.19 以及作业补充题 假设 $X\sim N(0,1)$, Y 以 0.5 的概率取值 ±1, 且 X,Y 相互独立。令 $Z=XY$, 证明

1. $Z\sim N(0,1)$;

2. X,Z 即不相关也不独立。

例 4.20, 作业 67 随机矩形的构造如下，底 X 选自 $(0,1)$ 上的均匀随机变量，生成完底部之后，取宽为 $[0,X]$ 上的均匀随机变量。利用全期望公式计算矩形的周长和面积。

解答：

由题意 $X \sim U(0,1)$ 和 $Y|X=x \sim U(0,x)$, 因此 $\mathbb{E}(Y|X=x)=x/2$, 所以

$$\begin{aligned} 2\mathbb{E}(X+Y) &= 2\mathbb{E}[\mathbb{E}(X+Y|X)] = 2\mathbb{E}(X+X/2) = 3/2; \\ \mathbb{E}(XY) &= \mathbb{E}[\mathbb{E}(XY|X)] = \mathbb{E}(X^2/2) = 1/6. \end{aligned}$$

例 4.21，作业 77 令 X,Y 的联合密度函数 $f(x,y)=e^{-y}, 0\le x\le y$.

1. 计算 $Cov(X,Y)$ 以及相关系数；
2. 计算 $\mathbb{E}(X|Y=y)$ 和 $\mathbb{E}(Y|X=x)$;
3. 推导随机变量 $\mathbb{E}(X|Y)$ 和 $\mathbb{E}(Y|X)$ 的密度函数。

解答：

1，由定义，需要先计算 X,Y,X^2,Y^2,XY 的期望

$$\begin{aligned} \mathbb{E}(X) &= \int_0^\infty \int_0^y xe^{-y}dxdy = 1; \\ \mathbb{E}(X^2) &= \int_0^\infty \int_0^y x^2e^{-y}dxdy = 2 \\ \mathbb{E}(Y) &= \int_0^\infty \int_0^y ye^{-y}dxdy = 2 \\ \mathbb{E}(Y^2) &= \int_0^\infty \int_0^y y^2e^{-y}dxdy = 6 \\ \mathbb{E}(XY) &= \int_0^\infty \int_0^y xye^{-y}dxdy = 3. \end{aligned}$$

则可得

$$\begin{aligned} Var(X) &= \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = 3; \\ Var(Y) &= \mathbb{E}(Y^2) - (\mathbb{E}(Y))^2 = 32; \\ Cov(X,Y) &= \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = 1. \end{aligned}$$

最后可得 X 和 Y 的相关系数为

$$Corr(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}} = \frac{1}{4\sqrt{6}}.$$

2, 由定义 $\mathbb{E}(X|Y=y)=\int_{-\infty}^{\infty}x^2 f_{X|Y}(x|y)dx$, 因此我们需要条件密度 $f_{X|Y}(x|y)$. 随机变量 Y 的边际密度函数为

$$
\begin{aligned}
f_Y(y) &= \int_{-\infty}^{\infty} f(x,y)dx \\
&= \begin{cases} \int_0^y e^{-y}dx, & 0<y; \\ 0, & otherwise. \end{cases} \\
&= \begin{cases} ye^{-y}, & 0<y; \\ 0, & otherwise. \end{cases}
\end{aligned}
$$

因此，在给定条件 $Y=y$ 时，X 的条件密度为

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

于是

$$
\begin{aligned}
\mathbb{E}(X|Y=y) &= \int_{-\infty}^{\infty} x f_{X|Y}(x|y)dx \\
&= \int_0^y x\frac{1}{y}dx \\
&= \frac{y}{2}
\end{aligned}
$$

同理，随机变量 X 的边际密度函数为

$$
\begin{aligned}
f_X(x) &= \int_{-\infty}^{\infty} f(x,y)dy \\
&= \begin{cases} \int_x^{\infty} e^{-y}dy, & 0<x; \\ 0, & otherwise. \end{cases} \\
&= \begin{cases} e^{-x}, & 0<x; \\ 0, & otherwise. \end{cases}
\end{aligned}
$$

因此，在给定条件 $X=x$ 时，Y 的条件密度为

$$
f_{Y|X}(y|x)=\begin{cases} e^{x-y}, & x<y; \\ 0, & otherwise. \end{cases}
$$

于是

$$
\begin{aligned}
\mathbb{E}(Y|X=x) &= \int_{-\infty}^{\infty} y f_{Y|X}(y|x)dy \\
&= \int_x^{\infty} ye^{x-y}dy \\
&= x+1
\end{aligned}
$$

3, 由第二问的结论，可知随机变量 $\mathbb{E}(X|Y)=Y/2$ 和 $\mathbb{E}(Y|X)=X+1$. 因此

$\mathbb{E}(X|Y)$ 的密度函数为

$$f_{\mathbb{E}(X|Y)}(y) = 2f_Y(2y) = \begin{cases} 4ye^{-2y}, & 0 < y; \\ 0, & otherwise. \end{cases}$$

和 $\mathbb{E}(Y|X)$ 的密度函数为

$$f_{\mathbb{E}(Y|X)}(x) = f_X(x-1) = \begin{cases} e^{1-x}, & 1 < x; \\ 0, & otherwise. \end{cases}$$

练习 4.22 令 $X_{(i)}$ 是区间 $(0,1)$ 上均匀随机变量 $X_1, X_2, \cdots, X_n$ 的次序统计量，$X_{(i)}$ 的密度函数为

$$f(x) = \frac{n!}{(i-1)!(n-i)!} x^{i-1}(1-x)^{n-i},\ 0 < x < 1.$$

1. 计算 $Var(X_{(i)})$;
2. 找出使得 $Var(X_{(i)})$ 达到最大值和最小值的 i.

练习 4.23 某箱装 100 件产品，其中一、二和三等品的件数分别为 80，10 和 10 件，现从中随机抽取一件，定义三个随机变量 X_1, X_2, X_3 如下

$$X_i = \begin{cases} 1, & \text{抽到第 i 等品;} \\ 0, & otherwise. \end{cases} \quad i = 1, 2, 3.$$

试求随机变量 X_1 和 X_2 的相关系数。

答案： -2/3.

练习 4.24 假设随机变量 (X, Y) 服从 $N(1, 0, 9, 16, -1/2)$, 设 $Z = \frac{X}{3} + \frac{Y}{2}$:

1. 求 Z 的 $\mathbb{E}(Z)$ 和 $D(Z)$;
2. 求 ρ_{XZ};
3. 问 Z, X 是否相互独立，为什么？

第 5 章
第五章：极限定理

5.1 大数律和中心极限定理

1. 依概率收敛：假设随机变量序列 $\xi_1, \xi_2, \cdots, \xi_n, \cdots$ 满足对任意 $\epsilon > 0$, 都有

$$\lim_{n\to\infty} \mathbb{P}(|\xi_n - \xi| > \epsilon) = 0$$

则称随机变量 ξ_n 依概率收敛到随机变量 ξ.

2. 伯努利大数律：假设 n_A 是 n 次独立重复实验中事件 A 发生的次数，且 $\mathbb{P}(A) = p$. 则对任意 $\epsilon > 0$, 有

$$\lim_{n\to\infty} \mathbb{P}(|\frac{n_A}{n} - p| \geq \epsilon) = 0.$$

3. **Chebyshev law of large number**：假设随机变量序列 $\{X_n\}_{n\in N^+}$ 为相互独立的随机变量列，且具有相同的期望和方差。记为

$$\mathbb{E}(X_i) = \mu,\ D(X_i) = \sigma^2, i = 1, 2, \cdots.$$

则对任意 $\epsilon > 0$, 有

$$\lim_{n\to\infty} \mathbb{P}(|\frac{1}{n}\sum_{1=1}^{n} X_i - \mu| > \epsilon) = 0$$

4. 辛钦大数律： 假设随机变量序列 $\{X_n\}_{n\in N^+}$ 相互独立，且具有相同的期望 μ. 则对任意 $\epsilon > 0$, 有

$$\lim_{n\to\infty} \mathbb{P}(|\frac{1}{n}\sum_{1=1}^{n} X_i - \mu| > \epsilon) = 0$$

5. 补充-强大数律： 假设随机变量序列 $\{X_n\}_{n\in N^+}$ 相互独立，且具有相同的期望 μ. 则以概率 1 的有

$$\lim_{n\to\infty} \frac{X_1 + \cdots + X_n}{n} = \mu.$$

6. 标准化：假设随机变量序列 $\{X_n\}_{n\in N^+}$ 为相互独立的随机变量列，且期望和方差都存在。记为 $\mathbb{E}(X_i) = \mu_i, D(X_i) = \sigma_i^2, i = 1, 2, \cdots$. 令 $Z_n = \frac{\sum_{i=1}^{n}(X_i - \mu_i)}{\sqrt{\sum_{i=1}^{n}\sigma_i^2}}$. 则

$$\mathbb{E}(Z_n) = 0,\ D(Z_n) = 1.$$

7. 定义：如果 Z_n 的分布函数

$$\lim_{n\to\infty} F_n(x) = \lim_{n\to\infty} \mathbb{P}(Z_n \le x) = \Phi(x),$$

则称随机变量序列 $\{X_n\}_{n\in N^+}$ 满足中心极限定理。

8. 中心极限定理：假设随机变量序列 $\{X_n\}_{n\in N^+}$ 独立同分布，且期望和方差都存在。记为

$$\mathbb{E}(X_i) = \mu,\ D(X_i) = \sigma^2, i = 1, 2, \cdots.$$

令 $Z_n = \frac{\sum_{i=1}^{n}(X_i-\mu)}{\sqrt{n}\sigma}$. 则 Z_n 的分布函数

$n \to \infty$

$$\lim_{n\to\infty} F_n(x) = \lim_{n\to\infty} \mathbb{P}(Z_n \le x) = \Phi(x),$$

9. 棣莫弗-拉普拉斯中心极限定理：假设随机变量序列 $\{\eta_n\}_{n\in N^+}$ 服从参数为 n, p 的二项随机变量。则 η_n 满足

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{\eta_n - np}{\sqrt{np(1-p)}} \le x\right) = \Phi(x),$$

例 5.1,Q1 设 $X_1, X_2, \cdots, X_n$ 是独立同分布的随机变量序列, 且 $\mathbb{E}(X_i) = \mu, D(X_i) = \sigma^2$, 那么 $\frac{1}{n}\sum_{i=1}^{n} X_i^2$ 依概率收敛于 $\sigma^2+\mu$ (请写出过程)

$D(X) = E(X^2) - E^2(X)$

解答：

由题意 $\mathbb{E}(X) = \mu$ 和 $D(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = \sigma^2$, 因此

$E(X^2) = \sigma^2 + \mu^2$

$$\mathbb{E}(X^2) = D(X) + [\mathbb{E}(X)]^2 = \sigma^2 + \mu^2.$$

利用辛钦大数律，$\mathbb{E}(X_i) = \mu, D(X_i) = \sigma^2$, 那么 $\frac{1}{n}\sum_{i=1}^{n} X_i^2$ 依概率收敛于 $\sigma^2 + \mu^2$.

例 5.2，Q2 设某工厂有 400 台同类机器，各台机器发生故障的概率都是 0.02, 各台机器工作是相互独立的，试求机器出故障的台数不小于 2 的概率（要求：用中心极限定理；可能用到的常数：$\Phi(2) = 0.9772; \Phi(2.5) = 0.9938; \Phi(3) = 0.9987.$）

X_i 为第 i 台机器故障 $\quad X_i = \begin{cases} 1 & 0.02 \\ 0 & 0.98 \end{cases}$

解答： $E(X_i) = 0.02 \quad D(X_i) = (1-0.02)^2 \times 0.02 + (0-0.02)^2 \times 0.98$

令随机变量 X 表示 400 台机器中出故障的台数，则可知 $X \sim B(400, 0.02)$. 由棣莫

$= 0.02 \times 0.98 = 0.0196$

$\sum_{i=1}^{400} X_i \sim N(8, 7.84)$

400

$P\{\sum_{i=1}^{400} x_i \geq 2\}$ $= P\{\frac{\sum_{i=1}^{400} x_i - 8}{\sqrt{7.84}} \geq \frac{2-8}{\sqrt{7.84}}\}$ $= 1-\Phi(-2.14)$ $= 1-(1-\Phi(2.14))$ $= \Phi(2.14)$

弗-拉普拉斯中心极限定理可知

$$\begin{aligned} \mathbb{P}(X \geq 2) &= \mathbb{P}(\frac{X-8}{\sqrt{7.84}} \geq \frac{-6}{\sqrt{7.84}}) \\ &\approx \Phi(2.14) \\ &\approx 0.9828 \end{aligned}$$

这里 Φ 是标准正态的分布函数。
或者

$$\begin{aligned} \mathbb{P}(X \geq 2) &= 1-\mathbb{P}(X \leq 1) \\ &\approx 1-\mathbb{P}(\frac{X-8}{\sqrt{7.84}} \leq \frac{-7}{\sqrt{7.84}}) \\ &= 1-\Phi(-2.5) \\ &\approx 0.9938 \end{aligned}$$

不同的计算思路误差在 0.01 左右。

例 5.3，Q3-B 为确定某个城市成年男子中吸烟者的比例 p, 任意调查 n 个成年男子，记其中吸烟人数为 m, 问 n 多大才能保证 m/n 与 p 的差异小于 0.01 的概率大于 0.95. (可能用到的常数 $\Phi(1.96)=0.975$.)

解答：
由题意 $m \sim B(n,p)$, 因此由棣莫弗-拉普拉斯中心极限定理可知

$$\begin{aligned} \mathbb{P}(|\frac{m}{n}-p| \leq 0,01) &= \mathbb{P}(|\frac{m-np}{n}| \leq 0.01) \\ &= \mathbb{P}(|\frac{m-np}{\sqrt{np(1-p)}\sqrt{\frac{n}{p(1-p)}}} \leq 0.01) \\ &= \mathbb{P}(|\frac{m-np}{\sqrt{np(1-p)}} \leq 0.01\sqrt{\frac{n}{p(1-p)}}) \\ &\approx 2\Phi(0.01\sqrt{\frac{n}{p(1-p)}})-1 > 0.95 \end{aligned}$$

上面最后一行不等式要求

$$\Phi(0.01\sqrt{\frac{n}{p(1-p)}}) > 0.975.$$

查表得

$$0.01\sqrt{\frac{n}{p(1-p)}} > 1.96, \quad or \quad n \geq 196^2 p(1-p).$$

如果还记得 $p(1-p) \le 1/4$, 可以给出一个下界 $n \ge 196^2/4 = 9604$.

例 5.4，Q3-A 一家有 500 间客房的大旅馆的每间客房有一台 2kw 空调机，若开房率为 0.8，需要多少 kw 的电力才能有 0.99 的可能保证有足够的电力使用空调机。

解答：

令 Y 为 500 间客房开房的房间数，可知 $Y \sim B(500, 0.8)$. 假设共有 k (KW) 的电力可供使用，由题意

$$\mathbb{P}(2Y \le k) = \mathbb{P}(Y \le k/2) \ge 0.99.$$

由棣莫弗-拉普拉斯中心极限定理可知

$$\begin{aligned} \mathbb{P}(Y \le k/2) &= \mathbb{P}(\frac{Y-400}{8\sqrt{5}} \le \frac{k/2-400}{4\sqrt{5}}) \\ &\approx \Phi(\frac{k/2-400}{4\sqrt{5}}) \\ &= \Phi(\frac{k-800}{8\sqrt{5}}) \end{aligned}$$

因此问题为

$$\Phi(\frac{k-800}{8\sqrt{5}}) \ge 0.99.$$

查表得 $\frac{k-800}{8\sqrt{5}} \ge 2.33$, 从中解得 $k \ge 841.68$, 取为 842KW 即可。这表明至少需要 842KW 才能以 0.99 的可能性保证空调通电。

第 6 章

第六章：数理统计-基本概念

1. 数理统计：它使用概率论和数学的方法，研究怎么收集（通过实验或者观察）带有随机误差的数据，并在设定的模型（成为统计模型）下，对这种数据进行分析（称为统计分析），以对所研究的问题作出推断（称为统计推断）。

2. 总体：研究对象的全体；个体：总体中的一个具体对象；样本：从研究对象中任取 n 个个体，观察他们的数量指标 $X_1, \cdots, X_n$, 称之为容量为 n 的样本；样本值：观测后，$x_1, x_2, \cdots, x_n$ 是一组具体数字。

3. 统计量：关于样本 $X_1, \cdots, X_n$ 的函数 $g(X_1, \cdots, X_n)$ 称之为统计量。注意，统计量本身是一个随机变量。

4.
   - χ^2 分布：设 $X_1, X_2, \cdots, X_n$ 是来自总体 $X \sim N(0,1)$ 的样本，令

     $$Y = X_1^2 + X_2^2 + \cdots + X_n^2$$

     则称 Y 服从自由度为 n 的 χ^2-分布。
   - χ^2-分布的密度函数

     $$f(y) = \begin{cases} \frac{1}{2^{n/2}\Gamma(n/2)} y^{n/2-1} e^{-y/2}, & y > 0; \\ 0, & y \leq 0. \end{cases}$$

   - χ^2 分布具有可加性：设 $U \sim \chi^2(n_1), V \sim \chi^2(n_2)$ 且相互独立，则 $U + V \sim \chi^2(n_1 + n_2)$;
   - 如果随机变量 $X \sim \chi^2(n)$, 则其期望和方差满足 $\mathbb{E}(X) = n, D(X) = 2n.$

5.
   - t 分布：设随机变量 $X \sim N(0,1), Y \sim \chi^2(n)$, 且 X, Y 相互独立，令

     $$Z = \frac{X}{\sqrt{Y/n}}$$

     则称 Z 服从自由度为 n 的 t-分布。
   - t-分布的密度函数

     $$f(y) = \frac{\Gamma[(n+1)/2]}{\sqrt{n\pi}\Gamma(n/2)}(1 + \frac{y^2}{n})^{-(n+1)/2}.$$

- t 分布的密度函数关于 y-轴对称；
- 如果随机变量 $X \sim t(n)$, 则其期望和方差满足 $\mathbb{E}(X) = 0(n > 1), D(X) = \frac{n}{n-2}(n > 2)$.

6.
- F 分布：设随机变量 $X \sim \chi^2(n_1), Y \sim \chi^2(n_2)$, 且 X, Y 相互独立，令

$$Z = \frac{X/n_1}{Y/n_2}$$

  则称 Z 服从自由度为 (n_1, n_2) 的 F-分布，记为 $Z \sim F(n_1, n_2)$.

- F-分布的密度函数

$$f(y) = \begin{cases} \frac{\Gamma[(n_1+n_2)/2]}{\Gamma(n_1/2)\Gamma(n_2/2)} n_1^{n_1/2} n_2^{n_2/2} \frac{x^{n_1/2-1}}{(n_1 x+n_2)^{(n_1+n_2)/2}}, & y > 0; \\ 0, & y \leq 0. \end{cases}$$

- 如果随机变量 $Z \sim F(n_1, n_2)$, 则随机变量 $\frac{1}{Z} \sim F(n_2, n_1)$；
- 如果随机变量 $X \sim F(n_1, n_2)$, 则其期望和方差满足 $\mathbb{E}(X) = \frac{n_2}{n_2-2}(n_2 > 2), D(X) = \frac{2n_2^2(n_1+n_2-2)}{n_1(n_2-2)^2(n_2-4)}(n_2 > 4)$.
- 如果 $X \sim t(n)$, 则 $X^2 \sim F(1, n)$.

7. α 分位数：常数 x_α 满足 $\mathbb{P}(X \leq x_\alpha) = \int_{-\infty}^{x_\alpha} f(x)dx = \alpha$;

8. F 分布的特殊性质：$F_\alpha(n_1, n_2) = \frac{1}{F_{1-\alpha}(n_2, n_1)}$.

9. 设 $X_1, X_2, \cdots, X_n$ 来自于总体 $X \sim N(\mu, \sigma^2)$, 则样本均值 $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$;

10. 设 $X_1, X_2, \cdots, X_n$ 来自于总体 $X \sim N(\mu, \sigma^2)$, 样本均值和样本方差分别为 $\bar{X}$ 和 S^2, 则 $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$;

11. 条件同上，则样本均值和样本方差相互独立且 $\frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$.

12. 设 $X_1, X_2, \cdots, X_{n_1}$ 是来自于总体 $X \sim N(\mu_1, \sigma_1^2)$ 的样本，$Y_1, Y_2, \cdots, Y_{n_2}$ 是来自于总体 $X \sim N(\mu_2, \sigma_2^2)$ 的样本，且两样本相互独立。此外，假设两样本均值和方差分别为 $\bar{X}, \bar{Y}, S_1^2, S_2^2$, 则

$$\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F(n_1-1, n_2-1).$$

13. 条件同上，则

$$\frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{S_\omega \sqrt{1/n_1 + 1/n_2}} \sim t(n_1 + n_2 - 2),$$

其中 $S_\omega^2 = \frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$.

例 6.1，Q12-1 从总体 $N(100,4)$ 中独立地进行两次抽样, 容量分别为 15 和 20, 那么这两个样本均值之差的绝对值大于 0.2 的概率是多少?

解答：
假设两个样本的样本均值分别为 $\bar{X}$ 和 $\bar{Y}$, 则可知 $\bar{X} \sim N(100, \frac{4}{15}), \bar{Y} \sim N(100, \frac{1}{5})$, 因此 $\bar{X} - \bar{Y} \sim N(0, \frac{7}{15})$. 可以得到

$$\begin{aligned}
\mathbb{P}(|\bar{X} - \bar{Y}| > o.2) &= \mathbb{P}(\frac{|\bar{X} - \bar{Y}|}{\sqrt{7/15}} > \frac{0.2}{\sqrt{7/15}}) \\
&= 1 - (2\Phi(\frac{0.2}{\sqrt{7/15}}) - 1) \\
&\approx 2 - 2\Phi(0.293) \\
&\approx 0.7718.
\end{aligned}$$

例 6.2，Q12-2 设总体 $X \sim N(\mu, \sigma^2), X_1, X_2, \cdots, X_n, X_{n+1}$ 为样本. 令

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n}, \quad S^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X})^2.$$

求 $\frac{X_{n+1} - \bar{X}}{S}\sqrt{\frac{n}{n+1}}$ 的分布。

解答：
由题意 $X_{n+1} \sim N(\mu, \sigma^2), \bar{X} \sim N(\mu, \frac{\sigma^2}{n})$, 因此 $X_{n+1} - \bar{X} \sim N(0, \frac{(n+1)\sigma^2}{n})$, 可得

$$\frac{X_{n+1} - \bar{X}}{\sqrt{(n+1)\sigma^2/n}} \sim N(0,1).$$

利用定理 2 的结论，可知 $\bar{X}, S^2$ 相互独立且 $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$, 因此可知

$$\frac{X_{n+1} - \bar{X}}{S}\sqrt{\frac{n}{n+1}} = \frac{\frac{X_{n+1} - \bar{X}}{\sqrt{(n+1)\sigma^2/n}}}{\sqrt{\frac{(n-1)S^2}{\sigma^2}/(n-1)}}$$

因此 $\frac{X_{n+1} - \bar{X}}{S}\sqrt{\frac{n}{n+1}} \sim t(n-1)$.

例 6.3,Q12-3A 设 $X_1, X_2, \cdots, X_9$ 和 $Y_1, Y_2, \cdots, Y_{16}$ 是分别来自总体 $N(0,4), N(1,9)$ 的两个样本, 且两个样本独立.. 求:

1. $\mathbb{P}(Y_1 < 0)$;
2. $\mathbb{P}(X_1 + Y_1 > 1)$;
3. $\bar{X} - \bar{Y}$ 服从什么分布。

解答：

1. 由题意，可知 $Y_1 \sim N(1, 9)$, 可得

$$\begin{aligned}\mathbb{P}(Y_1 < 0) &= \mathbb{P}(\frac{Y_1 - 1}{3} < -1/3) \\ &= \Phi(-1/3) \\ &\approx 0.3707.\end{aligned}$$

2. 由题意，可知 $X_1 \sim N(0, 4), Y_1 \sim N(1, 9)$, 因此 $X_1 + Y_1 \sim N(1, 13)$, 可得 $\mathbb{P}(X_1 + Y_1 > 1)$;

$$\begin{aligned}\mathbb{P}(X_1 + Y_1 > 1) &= \mathbb{P}(\frac{X_1 + Y_1 - 1}{13} > 0) \\ &= \Phi(0) \\ &\approx 0.5.\end{aligned}$$

3. 由定理 1 可得 $\bar{X} \sim N(0, \frac{4}{9}), \bar{Y} \sim N(1, \frac{9}{16})$, 因此

$$\bar{X} - \bar{Y} \sim N(-1, \frac{145}{144}).$$

例 6.4，Q12-4A 假设随机变量 X, Y 均服从标准正态分布，则下列判断正确的是 ( )

A: $X + Y$ 服从正态分布；

B: $X^2 + Y^2$ 服从 χ^2 分布；

C: X^2 和 Y^2 都服从 χ^2 分布；

D: $\frac{X^2}{Y^2}$ 服从 F 分布.

解答：

C，分析：由定义，标准正态的平方服从 $\chi^2(1)$, 因此 C 肯定成立。A, B, D 当 $Y=-X$ 时显然不成立。

例 6.5，Q12-3B 设 X_1, X_2, X_3 来自均值为 μ 方差为 σ^2 的总体. 求：

1. $D(X_1/2+X_2/3+X_3)$;
2. $\mathbb{E}(X_1+X_2+X_3)^2$.

解答：

1. 由题意 X_1, X_2, X_3 独立同分布，且均值为 μ 方差为 σ^2. 可知
$$D(X_1/2+X_2/3+X_3) = D(X_1)/4 + D(X_2)/9 + D(X_3) = \frac{49}{36}\sigma^2.$$
2. 由题意可知
$$\begin{aligned}\mathbb{E}(X_1+X_2+X_3) &= 3\mu \\ D(X_1+X_2+X_3) &= 3\sigma^2.\end{aligned}$$
因此
$$\mathbb{E}(X_1+X_2+X_3)^2 = D(X_1+X_2+X_3) + (\mathbb{E}(X_1+X_2+X_3))^2 = 9\mu^2+3\sigma^2.$$

例 6.6, Q12-4B 设 X_1, X_2, X_3, X_4 来自正态总体 $N(1,2)$ 的样本. 令 $Y = k[\sum_{i=1}^4 X_i - 4]^2$, 若要使 Y 服从 χ^2 分布，则 k 的取值和自由度（ ）

A: $k=\sqrt{1/8}$, 自由度为 4;

B: $k=1/8$, 自由度为 4;

C: $k=\sqrt{1/8}$, 自由度为 1;

D: $k = 1/8$, 自由度为 1.

解答：

C，分析：由题意 $\sum_{i=1}^{4} X_i - 4 \sim N(0,8)$, 因此选 C。

例 6.7，作业 Q8 证明：如果 X, Y 是参数 $\lambda = 1$ 的独立指数随机变量，那么 X/Y 服从 F 分布，同时指出自由度。

解答：

观察 Γ 分布的密度函数

$$f(y) = \begin{cases} \frac{1}{2^{n/2}\Gamma(n/2)} y^{n/2-1} e^{-y/2}, & y > 0; \\ 0, & y \le 0. \end{cases}$$

当 $n = 2$ 时，可知

$$f(y) = \begin{cases} \frac{1}{2\Gamma(1)} e^{-y/2}, & y > 0; \\ 0, & y \le 0. \end{cases}$$

由密度函数的性质（积分等于 1），可知 $\frac{1}{2\Gamma(1)} = 1/2$. 即 $\chi^2(2)$ 就是参数为 $1/2$ 的指数分布。此外，如果 $X \sim Exp(1)$, 则随机变量 $cX \sim Exp(1/c), c > 0$. 由此可知，$2X \sim \chi^2(2), 2Y \sim \chi^2(2)$, 由 F 分布定义可得

$$X/Y = \frac{2X/2}{2Y/2} \sim F(2,2).$$

练习 6.8 某工厂生产的灯泡使用寿命 $X \sim N(2250, 250^2)$, 现进行质量检查，方法如下：随机的抽取若干个灯泡，如果这些灯泡的平均寿命超过 2200 h，则认为该厂生产的灯泡合格，若要使检查通过的概率不低于 0.997，问至少要检查多少只灯泡。

练习 6.9 设总体服从参数为 λ 的指数分布，即分布密度为 $f(x) = \lambda e^{-\lambda x}, x > 0$. 求 $\mathbb{E}(\bar{X}), D(\bar{X}), \mathbb{E}(S^2)$.

练习 6.9

- 假设样本 $X_1, X_2, \cdots, X_6$ 来自总体 $N(0,1), Y = (X_1+X_2+X_3)^2+(X_4+X_5+X_6)^2$. 试确定常数 C 使得 CY 服从 χ^2 分布。
- 假设样本 $X_1, X_2, \cdots, X_5$ 来自总体 $N(0,1), Y = C\frac{X_1+X_2}{\sqrt{X_3^2+X_4^2+X_5^2}}$. 试确定常数 C 使得 Y 服从 t 分布。

例 6.10 假设随机变量 $X \sim F(n,n)$, 证明: $\mathbb{P}(X<1)=0.5$.

解答:

注意 F 分布的特殊性质: 如果 $Y \sim F(n,m)$, 则 $\frac{1}{Y} = F(m,n)$. 由此，可知 $1/X \sim F(n,n)$, 更进一步 $\mathbb{P}(X<1)=\mathbb{P}(1/X<1)=\mathbb{P}(X>1)$, 结合 $\mathbb{P}(X<1)+\mathbb{P}(X>1)=1$, 所以 $\mathbb{P}(X<1)=0.5$.

例 6.11，作业补充 3 假设 X_1, X_2 是来自于总体 $N(0,\sigma^2)$ 的样本，

1. 求 $\frac{(X_1+X_2)^2}{(X_1-X_2)^2}$ 的分布；
2. 求常数 k 使得 $\mathbb{P}(\frac{(X_1+X_2)^2}{(X_1-X_2)^2+(X_1+X_2)^2} > k) = 0.05$.

解答:

1. 由题意可知 (X_1+X_2, X_1-X_2) 服从二元正态，且满足

$$Cov(X_1-X_2, X_1+X_2) = \mathbb{E}(X_1^2) - \mathbb{E}(X_2^2) = 0,$$

因此 X_1-X_2, X_1+X_2 相互独立，均服从 $N(0,2\sigma^2)$. 可知 $\frac{(X_1-X_2)^2}{2\sigma^2} \overset{d}{=} \frac{(X_1+X_2)^2}{2\sigma^2} \sim \chi^2(1)$, 由 F 分布定义可知

$$\frac{(X_1+X_2)^2}{(X_1-X_2)^2} = \frac{(X_1+X_2)^2/2\sigma^2}{(X_1-X_2)^2/2\sigma^2} \sim F(1,1).$$

2. 假设 $F = \frac{(X_1+X_2)^2}{(X_1-X_2)^2}$, 则

$$\mathbb{P}(\frac{(X_1+X_2)^2}{(X_1-X_2)^2+(X_1+X_2)^2} > k) = \mathbb{P}(\frac{F}{1+F} > k)$$

上面我们将 $\frac{(X_1+X_2)^2}{(X_1-X_2)^2+(X_1+X_2)^2}$ 分子分母同时除以 $(X_1-X_2)^2$, 观察事件可得 $0 < k < 1$, 于是 $\frac{F}{1+F} > k$ 等价于不等式 $F > \frac{k}{1-k}$. 因此

$$\mathbb{P}(\frac{(X_1+X_2)^2}{(X_1-X_2)^2+(X_1+X_2)^2} > k) = \mathbb{P}(F > \frac{k}{1-k}).$$

查表可得 $\frac{k}{1-k} = F_{0.95}(1,1) = 161.45$, 这给出 $k = \frac{161.45}{1+161.45} = 0.9938$.

思考 6.12 设 $X_1, X_2, \cdots, X_{n_1}$ 是来自于总体 $X \sim N(\mu_1, \sigma^2)$ 的样本，$Y_1, Y_2, \cdots, Y_{n_2}$ 是来自于总体 $X \sim N(\mu_2, \sigma^2)$ 的样本，且两样本相互独立。c, d 是任意两个不为 0 的常数，试证明

$$\frac{c(\bar{X}-\mu_1)+d(\bar{Y}-\mu_2)}{S_\omega\sqrt{c^2/n_1+d^2/n_2}} \sim t(n_1+n_2-2),$$

其中 $S_\omega^2 = \frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$.

、解答：

由题意可知 $c(\bar{X}-\mu_1) \sim N(0, \frac{c^2\sigma^2}{n_1}$ 和 $d(\bar{Y}-\mu_2) \sim N(0, \frac{c^2\sigma^2}{n_2}$, 由此可得

$$c(\bar{X}-\mu_1)+d(\bar{Y}-\mu_2) \sim N(0, \frac{c^2\sigma^2}{n_1} + \frac{d^2\sigma^2}{n_2}).$$

利用定理 2，可知 $\frac{(n_1-1)S_1^2}{\sigma^2} \sim \chi^2(n_1-1)$ 和 $\frac{(n_2-1)S_2^2}{\sigma^2} \sim \chi^2(n_2-1)$, 利用样本之间的独立性，有

$$\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{\sigma^2} \sim \chi^2(n_1+n_2-2).$$

因此，我们有结论

$$\frac{c(\bar{X}-\mu_1)+d(\bar{Y}-\mu_2)}{S_\omega\sqrt{c^2/n_1+d^2/n_2}} = \frac{[c(\bar{X}-\mu_1)+d(\bar{Y}-\mu_2)]/\sqrt{\frac{c^2\sigma^2}{n_1}+\frac{d^2\sigma^2}{n_2}}}{\sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}},$$

思考 6.13 假设 $(X_i, Y_i), i = 1, 2, \cdots, n$ 是取自二维正态分布 $N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$ 的一个二维样本，记

$$\begin{aligned}
\bar{X} &= \frac{1}{n}\sum_{i=1}^n X_i, \bar{Y} = \frac{1}{n}\sum_{i=1}^n Y_i, \\
S_X^2 &= \frac{1}{n}\sum_{i=1}^n (X_i-\bar{X})^2, S_Y^2 = \frac{1}{n}\sum_{i=1}^n (Y_i-\bar{Y})^2, \\
\hat{\rho} &= \frac{\sum_{i=1}^n (X_i-\bar{X})(Y_i-\bar{Y})}{\sqrt{\sum_{i=1}^n (X_i-\bar{X})^2 \sum_{i=1}^n (Y_i-\bar{Y})^2}}.
\end{aligned}$$

试求统计量 $T = \sqrt{n-1}\frac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{\sqrt{S_X^2+S_Y^2-2\hat{\rho}S_XS_Y}}$ 的分布。

第 7 章
第七章：参数估计

7.1 点估计及其评价标准

1. 参数空间：假设总体 $X \sim F(x;\theta_1,\cdots,\theta_m)$, 其中函数形式已知，参数 $\theta = (\theta_1,\cdots,\theta_m)$ 的取值范围称为参数空间，记为 Θ.

2. 点估计：构造一个估计量 $\hat{\theta}(x_1,x_2,\cdots,x_n)$, 将其作为参数 θ 估计值。和样本的二重性一样，点估计应具有二重性:
   - 估计量 $\hat{\theta}(X_1,\cdots,X_n)$: 统计量的理论性质；
   - 估计值 $\hat{\theta}(x_1,\cdots,x_n)$: 统计推断。

3. 矩估计方法：假设总体 $X \sim F(x;\theta_1,\cdots,\theta_m)$, 其中未知参数为 $\theta = (\theta_1,\cdots,\theta_m)$，$X_1,X_2,\cdots,X_n$ 是来自该总体的样本。假设 $\mathbb{E}(X^k), k=1,2,\cdots,m$ 存在，由辛钦大数律 $\frac{1}{n}\sum_{i=1}^n X_i^k \approx \mathbb{E}(X^k) = \alpha_k$. 因此用 $\hat{\theta}_j = \hat{\theta}(A_1,A_2,\cdot,A_m), j=1,\cdots,m$. 作为参数 θ_j 的一个估计，称之为矩估计。

4. **Remark**：尽可能使用低阶矩；
   - 先求总体分布的理论矩，反解出未知参数；
   - 用样本矩代替理论矩作为未知参数的估计值。

5. **Fisher** 极大似然原理：极大似然原理的直观想法是，一个随机试验如有若干个可能的结果 A，B，C，... ，若在一次试验中，结果 A 出现了，那么可以认为实验条件对 A 的出现有利，也即出现的概率 $\mathbb{P}(A)$ 较大。极大似然原理的直观想法我们用下面例子说明。设甲箱中有 99 个白球，1 个黑球；乙箱中有 1 个白球．99 个黑球。现随机取出一箱，再从抽取的一箱中随机取出一球，结果是黑球，这一黑球从乙箱抽取的概率比从甲箱抽取的概率大得多，这时我们自然更多地相信这个黑球是取自乙箱的。

6. 极大似然估计: 假设 $X_1,X_2,\cdots,X_n$ 是总体 $X \sim f(x;\theta)$ 的样本，令

$$L(\theta) = L(\theta;X_1,\cdots,X_n) = \prod_{i=1}^n f(x_i;\theta)$$

称之为似然函数，如果存在 $\hat{\theta} = \hat{\theta}(X_1, \cdots, X_n)$ 使得

$$L(\hat{\theta}) = \max_{\theta \in \Theta} L(\theta; X_1, \cdots, X_n)$$

则称 $\hat{\theta}$ 为 θ 的极大似然估计，MLE.

7. 求极大似然估计的步骤：
   - 求似然函数：$L(\theta) = \prod_{i=1}^{n} f(x_i; \theta)$;
   - 求似然方程 $\frac{\partial}{\partial \theta_i} L(\theta) = 0$;
   - 解似然方程组，得到极大似然估计 $\hat{\theta}_i$.
8. 常用的方法和性质：
   - 因为对数函数是一个单调变换，因此使得对数似然函数：$\ell(\theta) = \ln L(\theta)$ 达到最大，也使得似然函数达到最大，因此经常使用对数似然求极大似然估计;
   - 极大似然估计的不变性：如果 $\hat{\theta}$ 是 θ 的极大似然估计，那么 $g(\hat{\theta})$ 是 $g(\theta)$ 的极大似然估计。
9. 无偏性
   如果统计量 $\hat{\theta} = \hat{\theta}(X_1, \cdots, X_n)$ 的期望存在，且满足

   $$\mathbb{E}(\hat{\theta}) = \theta,$$

   则称其为无偏估计。如果 $b_n(\hat{\theta})$ 收敛到 0，则称之为渐进无偏估计。
10. 有效性
    假设估计量 $\hat{\theta}_1, \hat{\theta}_2$ 都是参数 θ 的无偏估计，即 $\mathbb{E}(\hat{\theta}_1) = \mathbb{E}(\hat{\theta}_2) = \theta$. 如果有 $D(\hat{\theta}_1) < D(\hat{\theta}_2)$, 则称估计量 $\hat{\theta}_1$ 比 $\hat{\theta}_2$ 有效。Remark: 注意有效性是比较两个无偏估计。
11. 相合性如果对任意 $\varepsilon > 0$, 统计量 $\hat{\theta}_n = \hat{\theta}(X_1, \cdots, X_n)$ 满足

    $$\lim_{n \to \infty} \mathbb{P}(|\hat{\theta}_n - \theta| > \varepsilon) = 0$$

    则称 $\hat{\theta}_n$ 是参数 θ 的相合估计。
12. 均方误差统计量 $\hat{\theta}_n = \hat{\theta}(X_1, \cdots, X_n)$ 是参数 θ 的一个估计则称

    $$MSE(\hat{\theta}) = \mathbb{E}(\hat{\theta} - \theta)^2 = D(\hat{\theta}) + (\mathbb{E}(\hat{\theta}) - \theta)^2$$

    为 θ 的均方误差。
    - 均方误差是评价点估计好坏的比较一般的方法;

- 使得均方误差一致最小的估计量一般是不存在的，但可以用来比较估计的好坏；
- 无偏估计的均方误差就是估计量的方差。

例 7.1，Q13-1B 假设总体 X 的密度函数为 $f(x;\theta)=\frac{1}{2\theta}e^{-\frac{|x|}{\theta}}$. 已知 $X_1,X_2,\cdots,X_n$ 是来自该总体的一个样本.

1. 求 θ 的矩估计量 $\hat{\theta}_1$;
2. 求 θ 的极大似然估计量 $\hat{\theta}_2$.

解答：

1. 注意因为 $\frac{x}{2\theta}e^{-\frac{|x|}{\theta}}$ 是奇函数，所以
$$\mathbb{E}(X)=\int_{-\infty}^{\infty}xf(x;\theta)dx=0$$
也就是利用一阶矩不能得到参数 θ 的估计。因此需要计算二阶矩;
$$\begin{aligned}\mathbb{E}(X^2) &= \int_{-\infty}^{\infty}\frac{x^2}{2\theta}e^{-\frac{|x|}{\theta}}\\ &= \int_{0}^{\infty}\frac{x^2}{\theta}e^{-\frac{x}{\theta}}\\ &= 2\theta^2\end{aligned}$$
其中第二个等式利用了对称性。因此 θ 的矩估计为
$$\hat{\theta}_1=\sqrt{\frac{\sum_{i=1}^n X_i^2}{2n}}.$$
2. 不难写出似然函数
$$L(\theta)=(\frac{1}{2\theta})^n e^{-\frac{\sum_{i=1}^n|X_i|}{\theta}}.$$
因此对数似然函数为
$$\ell(\theta)=\ln L(\theta)=-n\ln(2\theta)-\frac{\sum_{i=1}^n|X_i|}{\theta}.$$
将对数似然对 θ 求导并使之等于 0 得到对数似然方程
$$\ell'(\theta))=-\frac{n}{\theta}+\frac{\sum_{i=1}^n|X_i|}{\theta^2}=0.$$

解得
$$\hat{\theta}_2 = \frac{\sum_{i=1}^n |X_i|}{n}.$$
而 $\ell''(\theta))|_{\hat{\theta}_2} = (\frac{n}{\theta^2} - \frac{2\sum_{i=1}^n |X_i|}{\theta^3})|_{\hat{\theta}_2} = -\frac{n^3}{(\sum_{i=1}^n |X_i|)^3} < 0$. 因此 $\hat{\theta}_2$ 是 θ 的极大似然估计。

例 7.2，Q13-2 设总体 $X \sim U(0,\theta)$, 其中 θ 是未知参数。$X_1, X_2, \cdots, X_n$ 为来自该总体的一个样本. 令
$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$$
为样本均值。求

1. 证明: $\hat{\theta} = 2\bar{X}$ 是 θ 的无偏估计和相合估计;
2. 求 θ 的极大似然估计，它是无偏估计吗? 它是相合估计吗? (提示: 利用 Chebyshev 不等式)

解答：

1. 证：因为总体 $X \sim U(0,\theta)$，所以 $\mathbb{E}(X) = \frac{\theta}{2}, D(X) = \frac{\theta^2}{12}$, 从而
$$\mathbb{E}(\bar{X}) = \frac{\theta}{2}, D(\bar{X}) = \frac{\theta^2}{12n}.$$
于是 $\mathbb{E}(\hat{\theta}) = 2\mathbb{E}(\bar{X}) = \theta$, 这说明 $\hat{\theta} = 2\bar{X}$ 是 θ 的无偏估计。更进一步，对任意 $\epsilon > 0$
$$\begin{aligned} \mathbb{P}(|\hat{\theta} - \theta| > \epsilon) &= \mathbb{P}(|\hat{\theta} - \mathbb{E}(\hat{\theta})| > \epsilon) \\ &\le \frac{D(\hat{\theta})}{\epsilon^2} = \frac{\theta^2}{3\epsilon^2 n} \end{aligned}$$
这里不等式部分利用了 Chebyshev 不等式。因此 $\lim_{n\to\infty} \mathbb{P}(|\hat{\theta} - \theta| > \epsilon) = 0$, 这说明 $\hat{\theta}$ 是 θ 的相合估计。

2. 似然函数为
$$L(\theta) = \begin{cases} (\frac{1}{\theta})^n, & X_{(n)} \le \theta; \\ 0, & otherwise. \end{cases}$$
显然 $L(\theta)$ 是 θ 的减函数，而 $X_{(n)} \le \theta$, 因此 θ 的极大似然估计为 $\hat{\theta}_1 = X_{(n)}$. 下证明 $\hat{\theta}_1$ 的无偏性和相合性：利用之前所学知识，可知 $\hat{\theta}_1$ 的密度函数为
$$f(x) = \frac{n}{\theta^n} x^{n-1}, \ \ 0 < x < \theta.$$

因此
$$\mathbb{E}(\hat{\theta}_1)=\int_0^\theta \frac{n}{\theta^n}x^n dx=\frac{n}{n+1}\theta$$
因此 $\hat{\theta}_1$ 不是无偏估计，但是一个渐进无偏估计。

同理，可得 $D(\hat{\theta}_1)=\frac{n\theta^2}{(n+2)(n+1)^2}$. 对任意 $\epsilon>0$, 令 $\epsilon_1=\epsilon-\frac{\theta}{n+1}$, 可知当 n 充分大的时候，$\epsilon_1>\epsilon/2$. 利用 Chebyshev 不等式，可得

$$\begin{aligned}
\mathbb{P}(|\hat{\theta}_1-\theta|>\epsilon) &\le \mathbb{P}(|\hat{\theta}_1-\mathbb{E}(\hat{\theta}_1)|>\epsilon_1)\\
&\le \frac{D(\hat{\theta}_1)}{\epsilon_1^2}\\
&\le \frac{D(4\hat{\theta}_1)}{\epsilon^2}
\end{aligned}$$

这里第一个不等式利用了
$$|\hat{\theta}_1-\theta|\le|\hat{\theta}_1-\mathbb{E}(\hat{\theta}_1)|+|\mathbb{E}(\hat{\theta}_1)-\theta|.$$

因此 $\lim_{n\to\infty}\mathbb{P}(|\hat{\theta}_1-\theta|>\epsilon)=0$, 这说明 $\hat{\theta}_1$ 是 θ 的相合估计。

例 7.3, Q13-1A 假设总体的密度为 $f(x;\theta)=\sqrt{\theta}x^{\sqrt{\theta}-1}, 0<x<1, \theta>0, X_1, X_2, \cdots, X_n$ 是来自该总体的样本，试求未知参数的极大似然估计。

解答：

不难写出似然函数
$$L(\theta)=(\sqrt{\theta})^n(\prod_{i=1}^n X_i)^{\sqrt{\theta}-1}.$$
因此对数似然函数为
$$\ell(\theta)=\ln L(\theta)=\frac{n}{2}\ln(\theta)+(\sqrt{\theta}-1)(\sum_{i=1}^n\ln X_i).$$
将对数似然对 θ 求导并使之等于 0 得到对数似然方程
$$\ell'(\theta))=\frac{n}{2\theta}+\frac{\sum_{i=1}^n\ln X_i}{2\sqrt{\theta}}=0.$$
解得
$$\hat{\theta}=(\frac{\sum_{i=1}^n\ln X_i}{n})^{-2}$$
而 $\ell''(\theta))|_{\hat{\theta}}=(-\frac{n}{2\theta^2}-\frac{2\sum_{i=1}^n\ln X_i}{4\theta^{3/2}})|_{\hat{\theta}}=-\frac{(\sum_{i=1}^n\ln X_i)^4}{4n^3}<0$. 因此 $\hat{\theta}$ 是 θ 的极大似然估计。

练习 7.4 假设总体的密度如下，$X_1, X_2, \cdots, X_n$ 是来自该总体的样本，试求未知参数的矩估计：

- $f(x;\theta) = \frac{2}{\theta^2}(\theta - x), 0 < x < \theta, \theta > 0;$
- $f(x;\theta) = (\theta+1)x^\theta, 0 < x < 1, \theta > 0;$
- $f(x;\theta) = \sqrt{\theta}x^{\sqrt{\theta}-1}, 0 < x < 1, \theta > 0;$
- $f(x;\theta,\mu) = \frac{1}{\theta}\exp(-\frac{x-\mu}{\theta}), \mu < x, \theta > 0;$

练习 7.5 假设总体的密度如下，$X_1, X_2, \cdots, X_n$ 是来自该总体的样本，试求未知参数的极大似然估计：

- $f(x;\theta) = \sqrt{\theta}x^{\sqrt{\theta}-1}, 0 < x < 1, \theta > 0;$
- $f(x;\theta) = \theta c^\theta x^{-(\theta+1)}, 0 < c < x, \theta > 1.$

练习 7.6 假设总体的密度如下，$X_1, X_2, \cdots, X_n$ 是来自该总体的样本，试求未知参数的极大似然估计：

- $f(x;\theta) = c\theta^c x^{-(c+1)}, 0 < \theta < x, c > 0;$
- $f(x;\theta,\mu) = \frac{1}{\theta}\exp(-\frac{x-\mu}{\theta}), \mu < x, \theta > 0;$
- $f(x;\theta) = \frac{1}{k\theta}, \theta < x < (k+1)\theta, \theta > 0;$

练习 7.7 假设总体的密度如下，$X_1, X_2, \cdots, X_n$ 是来自该总体的样本，试求未知参数的极大似然估计：

- $f(x;\theta) = \frac{1}{2\theta}e^{-|x|/\theta}, \theta > 0;$
- $f(x;\theta) = 1, \theta - 1/2 < x < \theta + 1/2;$
- $f(x;\theta_1,\theta_2) = \frac{1}{\theta_2-\theta_1}, \theta_1 < x < \theta_2.$

练习 7.8 设 X_1, X_2, X_3 是取自某总体的容量为 3 的样本，试证以下统计量都是该总体均值 μ 的无偏估计，在方差存在时指出哪个估计的有效性最差

1. $\hat{\mu}_1 = \frac{1}{2}X_1 + \frac{1}{3}X_2 + \frac{1}{6}X_3$;
2. $\hat{\mu}_2 = \frac{1}{3}X_1 + \frac{1}{3}X_2 + \frac{1}{3}X_3$;
3. $\hat{\mu}_3 = \frac{1}{6}X_1 + \frac{1}{6}X_2 + \frac{2}{3}X_3$;

练习 7.9 设 $\hat{\theta}$ 是参数 θ 的无偏估计，且有 $Var(\hat{\theta}) > 0$, 试证 $(\hat{\theta})^2$ 不是 θ^2 的无偏估计。

练习 7.10 设总体 $X \sim Exp(1/\theta)$, $X_1, X_2, \cdots, X_n$ 是来自该总体的样本，试证：$\bar{X}$ 和 $nX_{(1)}$ 都是 θ 的无偏估计，并比较其有效性。

7.2 区间估计

1. 点估计是用一个点（数）去估计未知参数，顾名思义，区间估计是用一个区间去估计未知参数。区间估计的好处是把可能的误差用醒目的形式标识出来。(J.Neymen)

2. 假设总体 $X \sim F(x;\theta)$, 对 $\forall 0 < \alpha < 1$, 如果存在两个统计量

$$\underline{\theta} = \underline{\theta}(X_1, X_2, \cdots, X_n), \bar{\theta} = \bar{\theta}(X_1, X_2, \cdots, X_n)$$

   使得对任意 $\theta \in \Theta$ 有

$$\mathbb{P}(\underline{\theta} < \theta < \bar{\theta}) = 1 - \alpha$$

   则称 $[\underline{\theta}, \bar{\theta}]$ 为置信水平为 $1 - \alpha$ 置信区间。

3. 枢轴变量法 **(Pivot variable):**
   寻找区间估计的一般步骤

   (a) 找一个和未知参数 $g(\theta)$ 有关的统计量 T, 一般选取其良好的点估计;
   (b) 试找出统计量 T 和 $g(\theta)$ 的某一函数 $S(T, g(\theta))$, 其分布函数 F 和未知参数 θ 无关，这个函数 S 就成为枢轴变量;

(c) 对任意常数 $a<b$, 不等式 $a<S(T,g(\theta))<b$ 要能够等价的写成 $A<g(\theta)<B$ 的形式（A,B 与 θ 无关）；

(d) 把 a,b 取成 F 的分位点 $\omega_{\alpha/2},\omega_{1-\alpha/2}$, 反解出相应的 A,B, 区间 $[a,b]$ 就是置信水平为 $1-\alpha$ 的置信区间。

4. 单个正态总体均值的置信区间:
假设样本 $X_1,X_2,\cdots,X_n$ 为总体 $N(\mu,\sigma^2)$, 求 μ 置信水平为 $1-\alpha$ 的置信区间。

(a) 如果 σ^2 已知，可知枢轴变量 $\frac{\bar{X}-\mu}{\sqrt{\sigma^2/n}}\sim N(0,1)$, 则可知 μ 的区间估计为 $[\bar{X}-\sigma u_{1-\alpha/2}/\sqrt{n},\bar{X}+\sigma u_{1-\alpha/2}/\sqrt{n}]$;

(b) 如果 σ^2 未知，可知枢轴变量 $\frac{\bar{X}-\mu}{S/\sqrt{n}}\sim t(n-1)$, 则可知 μ 的区间估计为 $[\bar{X}-\frac{S}{\sqrt{n}}t_{1-\alpha/2}(n-1),\bar{X}+\frac{S}{\sqrt{n}}t_{1-\alpha/2}(n-1)]$;

5. 单个正态总体方差的置信区间：
假设样本 $X_1,X_2,\cdots,X_n$ 为总体 $N(\mu,\sigma^2)$, 求 σ^2 置信水平为 $1-\alpha$ 的置信区间。由之前的知识可知 S^2 是 σ^2 的无偏，相合估计。

(a) 可以构造枢轴变量 $\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$, 则可知 σ^2 的区间估计为 $[\frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)},\frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)}]$;

(b) $\frac{(n-1)S^2}{\chi^2(n-1)}$ 只是为了直观表示，没有什么意义！

6. 两正态总体均值差的置信区间：
假设样本 $X_1,X_2,\cdots,X_{n_1}$ 为总体 $N(\mu_1,\sigma_1^2)$, 样本 $Y_1,Y_2,\cdots,Y_{n_2}$ 为总体 $N(\mu_2,\sigma_2^2)$. 求 $\mu_1-\mu_2$ 置信水平为 $1-\alpha$ 的置信区间。由之前的知识可知 $\bar{X}-\bar{Y}$ 是 $\mu_1-\mu_2$ 的无偏，相合估计。

(a) 如果方差 σ_1^2,σ_2^2 已知，可以构造枢轴变量 $\frac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$, 则可知 $\mu_1-\mu_2$ 的区间估计为

$$[\bar{X}-\bar{Y}-u_{1-\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}},\bar{X}-\bar{Y}-u_{1-\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}].$$

(b) 如果方差 $\sigma_1^2=\sigma_2^2=\sigma^2$ 未知，这时 S_1^2,S_2^2 以及他们的加权平均 $S_\omega^2=\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$ 都是 σ^2 无偏估计，利用定理五可以构造枢轴变量 $\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{S_\omega\sqrt{1/n_1+1/n_2}}\sim t(n_1+n_2-2)$, 则可知 $\mu_1-\mu_2$ 的区间估计为

$$[(\bar{X}-\bar{Y})\pm t_{1-\alpha/2}S_\omega\sqrt{1/n_1+1/n_2}]$$

7. 两正态总体方差比的置信区间：
假设样本 $X_1,X_2,\cdots,X_{n_1}$ 为总体 $N(\mu_1,\sigma_1^2)$, 样本 $Y_1,Y_2,\cdots,Y_{n_2}$ 为总体 $N(\mu_2,\sigma_2^2)$. 求 $\frac{\sigma_1^2}{\sigma_2^2}$ 置信水平为 $1-\alpha$ 的置信区间。

(a) 由于 S_1^2, S_2^2 都是他们各自方差 σ_1^2, σ_2^2 无偏估计，利用定理五可以构造枢轴变量 $\frac{S_1^2/S_2^2}{\sigma_1^2/\sigma_2^2}$, 则可知 $\frac{\sigma_1^2}{\sigma_2^2}$ 的区间估计为

$$[\frac{S_1^2}{S_2^2}\frac{1}{F_{1-\alpha/2}(n_1-1,n_2-1)}, \frac{S_1^2}{S_2^2}\frac{1}{F_{\alpha/2}(n_1-1,n_2-1)}]$$

8. 大样本法:
大样本法就是利用极限分布，主要是中心及极限定理，建立枢轴变量，然后利用类似的步骤构造区间估计。

9. 单侧区间估计:
假设总体 $X \sim F(x;\theta)$, 对 $\forall 0 < \alpha < 1$, 如果存在统计量

$$\underline{\theta} = \underline{\theta}(X_1, X_2, \cdots, X_n), \bar{\theta} = \bar{\theta}(X_1, X_2, \cdots, X_n)$$

使得对任意 $\theta \in \Theta$ 有

$$\mathbb{P}(\underline{\theta} < \theta) = 1 - \alpha$$

则称其为置信水平为 $1-\alpha$ 单侧置信区间，称 $\underline{\theta}$ 为单侧置信下限。

例 7.11 关于置信水平为 95% 的置信区间下列说法正确的是:

1. A、以 95% 的概率包含总体均值;
2. B、有 5% 的可能性包含总体均值;
3. C、绝对包含总体均值;
4. D、绝对包含总体均值或绝对不包含总体均值.

解答:
答案选 D。因为在经典统计的范畴内，参数是一个确定的数，没有随机性，所以根据样本构造的区间估计要么包含参数真值，要么不包含参数真值。

例 7.12 下列说法正确的是:

1. A、95% 的置信区间将以 95% 的概率包含总体参数;

2. B、当样本量不变时，置信水平越大得到的置信区间就越窄;
3. C、当置信水平不变时，样本量越大得到的置信区间就越窄;
4. D、当置信水平不变时，样本量越大得到的置信区间就越宽.

解答：
答案选 C。

例 7.13 95% 的置信水平是指：

1. A、总体参数落在一个特定的样本所构造的区间内的概率为 95%;
2. B、用同样的方法构造的总体参数的多个区间中，包含总体参数的区间的比例为 95%;
3. C、总体参数落在一个特定的样本所构造的区间内的概率为 5%;
4. D、用同样的方法构造的总体参数的多个区间中，包含总体参数的区间的比例为 5%.

解答：
答案选 B。

例 7.14 在其他条件不变的情况下，提高抽样估计的可靠程度，其精度将（ ）
（A）增加 （B）不变 （C）减少 （D）以上都对
解答：
答案选 C。

例 7.15 在对正态总体均值进行区间估计时，如果方差已知，在其他条件不变的情况下，要使置信区间的宽度缩小一半，样本量应增加（ ）
（A）一倍 （B）两倍 （C）三倍 （D）四倍
解答：
答案选 C。

例 7.16 0.5.0, 1.25, 0.80, 2.00 是取自总体 X 的样本，已知 $Y = \ln X$ 服从正态分布 $N(\mu, 1)$.

1. 求 μ 的置信水平为 95% 的置信区间；

2. 求 X 的数学期望的置信水平为 95% 的置信区间。

解答：

1. 将数据进行对数变换，得到 $Y = \ln X$ 的样本值为

$$-0.6931, 0.2231, -0.2231, 0.6931.$$

它可以看做来自于正态总体 $N(\mu, 1)$, 其样本均值 $\bar{y} = 0$, 由于方差 $\sigma^2 = 1$ 已知，因此 μ 的置信水平为 95% 的置信区间为

$$[\bar{y} - u_{1-\alpha/2}/\sqrt{n}, \bar{y} + u_{1-\alpha/2}/\sqrt{n}] = [-0.9800, 0.9800].$$

2. 由于 $\mathbb{E}(X) = e^{\mu+1/2}$ 是 μ 的严格增函数，利用上一问的结果，可以算得 X 的数学期望的置信水平为 95% 的置信区间为

$$[e^{-0.98+0.5}, e^{0.98+0.5}] = [0.6188, 4.3929].$$

例 7.17 设总体 $X \sim N(\mu, \sigma^2)$, σ^2 已知，试问样本容量 n 多大才能保证 μ 的置信水平为 95% 的置信区间的长度不大于 k. 解答：
已知 μ 的置信水平为 95% 的置信区间为

$$[\bar{x} - u_{0.975}\sigma/\sqrt{n}, \bar{x} + u_{0.975}\sigma/\sqrt{n}]$$

其区间长度为 $2u_{0.975}\sigma/\sqrt{n}$, 若要使 $2u_{0.975}\sigma/\sqrt{n} \leq k$, 只需

$$n \geq (2/k)^2\sigma^2 u_{0.975}^2 = (\frac{3.92\sigma}{k})^2.$$

练习 7.18 用一个仪表测量某一物理量 9 次，得到样本平均值 $\bar{x} = 56.32$, 样本标准差为 $s = 0.22$.

1. 测量标准差 σ 反映了仪表的精度，试求 σ 置信水平为 0.95 的置信区间；

2. 求该物理量真值置信水平为 0.99 的置信区间。

练习 7.19 设从总体 $X \sin N(\mu_1, \sigma_1^2)$ 和总体 $Y \sim N(\mu_2, \sigma_2^2)$ 抽取容量为 $n_1 = 10, n_2 = 15$ 的样本，可以计算得 $\bar{x} = 82, s_x^2 = 56.5, \bar{y} = 76, s_y^2 = 52.4$.

1. 若已知 $\sigma_1^2 = 64, \sigma_2^2 = 49$, 试求 $\mu_1 - \mu_2$ 置信水平为 95% 的置信区间;

2. 若 $\sigma_1^2 = \sigma_2^2$, 试求 $\mu_1 - \mu_2$ 置信水平为 95% 的置信区间;

3. 试求 σ_1^2/σ_2^2 置信水平为 95% 的置信区间;

例 7.20 设 $X_1, X_2, \cdots, X_n$ 是来自 Poisson 分布 $P(\lambda)$ 的样本，试求当 n 足够大时，λ 置信水平为 $1-\alpha$ 的置信区间。**解答：**
由中心极限定理可知，当 n 比较大时，样本均值 $\bar{X} \stackrel{\cdot}{\sim} N(\lambda, \frac{\lambda}{n})$. 因而 $\frac{\bar{X}-\lambda}{\sqrt{\lambda/n}} \stackrel{\cdot}{\sim} N(0,1)$, 分布与 λ 无关，因此可以将 $\frac{\bar{X}-\lambda}{\sqrt{\lambda/n}}$ 作为枢轴变量。对于给定的 α, 利用标准正态的 $1-\alpha/2$ 可得

$$\mathbb{P}(|\frac{\bar{X}-\lambda}{\sqrt{\lambda/n}}| \le u_{1-\alpha/2}) = 1-\alpha,$$

这个事件等价于 $(\bar{X}-\lambda)^2 \le u_{1-\alpha/2}^2 \lambda/n$, 因而可得

$$\lambda^2 - (\bar{X} + \frac{1}{n}u_{1-\alpha/2}^2)\lambda + \bar{X}^2 \le 0$$

这是一个向上开口的抛物线，其判别式

$$(\bar{X} + \frac{1}{n}u_{1-\alpha/2}^2)^2 - 4\bar{X}^2 = \frac{4\bar{X}u_{1-\alpha/2}^2}{n} + \frac{u_{1-\alpha/2}^2}{n}^2 \ge 0,$$

因此这个抛物线和 λ 轴有两个交点 λ_1, λ_2, 满足

$$\mathbb{P}(\lambda_1 < \lambda < \lambda_2) = 1-\alpha.$$

λ_1, λ_2 的具体值自己求。

例 7.20 设 $X_1, X_2, \cdots, X_n$ 是来自指数分布 $Exp(\lambda)$ 的样本，试求 λ 置信水平为 $1-\alpha$ 的置信区间。**解答：**
首先，如果 $X \sim Exp(\lambda)$, 则 $cX \sim Exp(\frac{\lambda}{c})$. 由此，可知 $2\lambda X \sim Exp(1/2)$. 而参数为 $1/2$ 的指数分布恰好是 $\chi^2(2)$. 利用卡方分布的可加性，可知

$$2\lambda \sum_{i=1}^{n} X_i \sim \chi^2(2n).$$

因而可以选取 $2\lambda \sum_{i=1}^{n} X_i$ 作为枢轴变量，相应的置信区间为

$$[\frac{\chi_{\alpha/2}^2(2n)}{\sum_{i=1}^{n} X_i}, \frac{\chi_{1-\alpha/2}^2(2n)}{\sum_{i=1}^{n} X_i}].$$

练习 7.21 假设人体的身高服从正态分布，今抽测甲乙两地区 18-25 岁女青年的身高数据如下：甲地区抽取 10 名，样本均值为 1.64m，样本标准差为 0.2m。乙地区抽取 10 名，样本均值为 1.62m，样本标准差为 0.4m。求：

1. 两方差比的置信水平为 0.95 的置信区间；
2. 如果 $\sigma_1^2 = \sigma_2^2 = 1$, 两均值差的置信水平为 0.95 的置信区间；

练习 7.22 假为防止出厂产品缺斤少两，某厂质检人员从当天产品中随机抽取 12 包过称，称得重量（以 g 为单位）分别为：9.9, 10.1, 10.3, 10.4, 9.7, 9.8, 10.1, 10.0, 9.8, 10.3。假定重量服从正态分布，试以此数据对该产品的平均重量求置信水平为 95% 的区间估计。

练习 7.23 为估计某台光谱仪测量材料中金属含量的误差，特置备了 5 个金属试块，其成分、金属含量、均匀性都有差别，设每个试块的测量值都服从正态分布，现对每个试块重复测量 6 次，计算得标准差分别为 $s_1 = 0.09, s_2 = 0.11, s_3 = 0.14, s_4 = 0.10, s_5 = 0.11$, 求 σ 置信水平为 0.95 的置信区间。

第 8 章
第八章：假设检验

8.1 概述及正态总体的假设检验

1. 统计假设：参数空间 $\Theta = \{\theta\}$ 的非空子集或者关于参数 θ 的命题，称为统计假设，简称假设：

   (a) 原假设：根据需要而设立的假设，常记为 $H_0 : \theta \in \Theta_0$;

   (b) 备择假设：在原假设被拒绝后而采用的假设，常记为 $H_1 : \theta \in \Theta_1$;

   Remark: 要求 $\Theta_1 \cap \Theta_1 = \emptyset$, 即原假设 H_0 和备择假设 H_1 不能有公共参数。

2. 检验：利用所得到的样本，对原假设 $H_0 : \theta \in \Theta_0$ 作出判断的法则称为检验法则，简称为检验。检验有两个结果：

   (a) "原假设不正确"：称为拒绝原假设，或者检验显著；

   (b) "原假设正确"：称为接受原假设，或者检验不显著；

   Remark: 接受原假设并不意味着原假设正确，而是在现有的证据下，没有理由作出原假设不正确的判断。

3. 检验问题：由原假设 H_0 和备择假设 H_1 组成的需要做出判断的问题称为假设检验问题。关于参数 θ 的检验问题有如下三个，其中 θ_0 是已知常数：

   (a) $H_0 : \theta \leq \theta_0;\ vs\ H_1 : \theta > \theta_0$;

   (b) $H_0 : \theta \geq \theta_0;\ vs\ H_1 : \theta < \theta_0$;

   (c) $H_0 : \theta = \theta_0;\ vs\ H_1 : \theta \neq \theta_0$;

   Remark: 前面两个检验问题称为单侧检验问题，第三个是一个双侧检验问题。

4. 两类错误及其发生的概率：

   (a) 第一类错误： 原假设 H_0 正确，但被拒绝，这类错误一般称为第一类错误，其发生的概率称为第一类错误发生的概率或者拒真概率，记为 α;

(b) 第二类错误：原假设 H_1 正确，但被接受，这类错误一般称为第二类错误，其发生的概率称为第二类错误发生的概率或者受伪概率，记为 β;

5. 检验问题的基本步骤：

   (a) 建立假设，根据要求建立原假设 H_0 和备择假设 H_1;

   (b) 选择检验统计量，给出拒绝域 W 的形式：

      i. 用于对原假设 H_0 作出判断的统计量称为检验统计量；

      ii. 使原假设被拒绝的样本观察值所在的区域称为拒绝域，常用 W 表示；

      iii. 一个拒绝域对应着一个检验法则，反之，一个检验法则也对应着一个拒绝域。

   (c) 选择一个显著性水平 α, 只控制第一类错误的概率不超过 α 的检验一般称为显著性检验，α 一般取为 0.05，0.10 等。

   (d) 给出拒绝域，即满足 $\mathbb{P}(W)=\alpha$ 的区域。

   (e) 作出判断：

      i. 如果 $(X_1,X_2,\cdots,X_n)\in W$, 则拒绝 H_0, 称为接受 H_1;

      ii. 如果 $(X_1,X_2,\cdots,X_n)\in \bar{W}$, 则接受 H_0.

6. 正态总体的假设检验：

   (a) 单个正态 $N(\mu,\sigma^2)$ 均值 μ 的假设检验，显然样本均值是总体均值十分好的估计，:

      i. 如果总体的方差 σ^2 已知, 此时可以构造检验统计量 $\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$：三个检验问题的拒绝域分别为 $\{\bar{X}>\mu_0+\sigma/\sqrt{n}u_{1-\alpha}\},\{\bar{X}<\mu_0-\sigma/\sqrt{n}u_{1-\alpha}\},\{\bar{X}<\mu_0-\sigma/\sqrt{n}u_{1-\alpha/2}\}\cup\{\bar{X}>\mu_0+\sigma/\sqrt{n}u_{1-\alpha/2}\}$;

      ii. 如果总体的方差 σ^2 未知，此时可以构造检验统计量 $\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\sim t(n-1)$: 三个检验问题的拒绝域分别为 $\{\bar{X}>\mu_0+S/\sqrt{n}t_{1-\alpha}(n-1)\},\{\bar{X}<\mu_0-S/\sqrt{n}t_{1-\alpha}(n-1)\},\{\bar{X}<\mu_0-S/\sqrt{n}t_{1-\alpha/2}(n-1)\}\cup\{\bar{X}>\mu_0+S/\sqrt{n}t_{1-\alpha/2}(n-1)\}$;

   (b) 单个正态总体 $N(\mu,\sigma^2)$ 方差 σ^2 的检验，样本方差是总体方差一个十分好的估计量，可以构造检验统计量 $\frac{(n-1)S^2}{\sigma_0^2}$: 三个检验问题的拒绝域分别为 $\{S^2>\frac{\sigma_0^2\chi^2_{1-\alpha}(n-1)}{n-1}\},\{S^2<\frac{\sigma_0^2\chi^2_{\alpha}(n-1)}{n-1}\},\{S^2>\frac{\sigma_0^2\chi^2_{1-\alpha/2}(n-1)}{n-1}\}\cup\{S^2<\frac{\sigma_0^2\chi^2_{\alpha/2}(n-1)}{n-1}\}$.

   (c) 两个正态 $N(\mu_1,\sigma^2),N(\mu_2,\sigma^2)$ 均值差 $\mu_1-\mu_2$ 的假设检验，显然样本均值是总体均值十分好的估计，:

      i. 如果总体的方差 σ_1^2,σ_2^2 已知, 此时可以构造检验统计量 $\frac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{\sqrt{\sigma_1^2/+n_1\sigma_2^2/n_2}}$：三个检验问题的拒绝域分别为 $\{\bar{X}-\bar{Y}>(\mu_1-\mu_2)+\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}u_{1-\alpha}\},\{\bar{X}-$

$\bar{Y} < (\mu_1-\mu_2)-\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}u_{1-\alpha}\}, \{\bar{X}-\bar{Y} < (\mu_1-\mu_2)-\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}u_{1-\alpha/2}\}\cup$ $\{\bar{X}-\bar{Y} > (\mu_1-\mu_2)+\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}u_{1-\alpha/2}\}$;

ii. 如果总体的方差 σ_1^2, σ_2^2 未知，此时可以构造检验统计量 $\frac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{S_\omega\sqrt{1/n_1+1/n_2}}$：三个检验问题的拒绝域分别为 $\{\bar{X}-\bar{Y} > (\mu_1-\mu_2)+S_\omega\sqrt{1/n_1+1/n_2}t_{1-\alpha}(n_1+n_2-2)\}, \{\bar{X}-\bar{Y} < (\mu_1-\mu_2)-S_\omega\sqrt{1/n_1+1/n_2}t_{1-\alpha}(n_1+n_2-2\}, \{\bar{X}-\bar{Y} < (\mu_1-\mu_2)-S_\omega\sqrt{1/n_1+1/n_2}t_{1-\alpha/2}(n_1+n_2-2)\}\cup\{\bar{X}-\bar{Y} > (\mu_1-\mu_2)+S_\omega\sqrt{1/n_1+1/n_2}t_{1-\alpha/2}(n+m-2)\}$;

(d) 两个正态 $N(\mu_1,\sigma_1^2), N(\mu_2,\sigma_2^2)$ 方差的假设检验，显然方差比 S_X^2/S_Y^2 是一个十分好的检验统计量：三个检验的拒绝域分别为 $\{S_X^2/S_Y^2 > F_{1-\alpha}(n_1-1,n_2-1)\}, \{S_X^2/S_Y^2 < F_\alpha(n_1-1,n_2-1)\}, \{S_X^2/S_Y^2 > F_{1-\alpha/2}(n_1-1,n_2-1)\}\cup\{S_X^2/S_Y^2 < F_{\alpha/2}(n_1-1,n_2-1)\}$.

7. 指数分布均值 θ 的假设检验：$\frac{2n\bar{X}}{\theta_0}$ 是一个很好的检验统计量，三个检验问题的拒绝域为 $\{\frac{2n\bar{X}}{\theta_0} > \chi^2_{1-\alpha}(2n)\}, \{\frac{2n\bar{X}}{\theta_0} < \chi^2_{\alpha}(2n)\}, \{\frac{2n\bar{X}}{\theta_0} > \chi^2_{1-\alpha/2}(2n)\}\cup\{\frac{2n\bar{X}}{\theta_0} < \chi^2_{\alpha/2}(2n)\}$.

8. 大样本法：检验统计量 $\frac{\sqrt{n}(\bar{X}-p_0)}{\sqrt{p_0(1-p_0)}}$.

9. 假设检验和区间估计之间的关系：

(a) 双边检验问题的接受域 $\bar{W}$（拒绝域的补集）是置信水平为 $1-\alpha$ 的置信区间；

(b) 单边检验问题 $H_0:\theta\le\theta_0\ vs\ H_1:\theta>\theta_0$ 的接受域 $\bar{W}$（拒绝域的补集）可以得到置信水平为 $1-\alpha$ 的置信上限；

(c) 单边检验问题 $H_0:\theta\ge\theta_0\ vs\ H_1:\theta<\theta_0$ 的接受域 $\bar{W}$（拒绝域的补集）可以得到置信水平为 $1-\alpha$ 的置信下限；

基础概念：

1. 研究者想收集证据予以支持的假设通常设为（ ）
A. 原假设 B. 备择假设 C. 合理假设 D. 正常假设
答案：备择假设。

2. 在假设检验中，原假设和备择假设（ ）
A. 都有可能成立 B. 都有可能不成立 C. 只有一个成立而且必有一个成立 D. 原假设一定成立，备择假设不一定成立
答案：C，真实模型要么服从原假设，要么服从备择假设。

3. 在假设检验中，第 类错误是指（ ）
A. 当原假设正确时拒绝原假设 B. 当原假设错误时拒绝原假设 C. 当备择假设正

确时未拒绝备择假设 D. 当备择假设不正确时拒绝备择假设
答案：A，定义!

4. 在假设检验中，不拒绝原假设意味着（ ）。
A. 原假设肯定是正确的 B. 原假设肯定是错误的 C. 没有证据证明原假设是正确的 D. 没有证据证明原假设是错误的
答案：D

5. 若一项假设规定显著性水平为 $\alpha = 0.05$, 下面的表述哪一个是正确的（）
A. 接受 H_0 时的可靠性为 95% B. 接受 H_1 时的可靠性为 95% C. H_0 为假时被接受的概率为 5% D. H_1 为真时被拒绝的概率为 5%
答案：B

6. 进行假设检验时，在样本量一定的条件下，犯第一类错误的概率减小，犯第二类错误的概率就会（）
A. 减小 B. 增大 C. 不变 D. 不确定
答案：B

7. 在一次假设检验中当显著性水平 $\alpha = 0.01$, 原假设被拒绝时，则用 $\alpha = 0.05$ 时，()
A. 原假设一定会被拒绝 B. 原假设一定不会被拒绝 C. 需要重新检验 D. 有可能拒绝原假设答案：A

8. 某厂生产的化纤纤度服从正态分布，纤维纤度的标准均值为 1.40。某天测得 25 根纤维的纤度的均值为 $\bar{x} = 1.39$, 检验与原来设计的标准均值相比是否有所下降，要求的显著性水平为 $\alpha = 0.05$, 则下列正确的假设形式是（）
A. $H_0 : \mu = 1.40\ vs\ H_1 : v \neq 1.40$; B. $H_0 : \mu \leq 1.40\ vs\ H_1 : v > 1.40$; C. $H_0 : \mu < 1.40\ vs\ H_1 : v \geq 1.40$; D. $H_0 : \mu \geq 1.40\ vs\ H_1 : v < 1.40$;
答案 D

9. . 当原假设正确而被拒绝时，所犯的错误为__;c 当备择假设正确而未拒绝原假设时，我们所犯的错误为__。只有在拒绝原假设时我们才可能犯第__ 类错误。只有在接受原假设时我们才可能犯第__ 类错误。

10. 有个研究者猜测，某贫困地区失学儿童中女孩数是男孩数的 3 倍以上（即失学男孩数不足失学女孩数的 1/3）。为了对他的这一猜测进行检验，拟随机抽取 50 个失学儿童构成样本。试问：这里要检验的参数是__，原假设和备择假设分别__，采用的检验统计量形式为__。

例 8.1 假设样本 $X_1, X_2, \cdots, X_n$ 来自于总体 $N(\mu, 1)$，考虑如下的检验问题

$$H_0 : \mu = 2;\ vs\ \mu = 3.$$

若检验由拒绝域 $W := \{\bar{X} > 2.6\}$ 确定。

1. 当 $n = 20$ 时，求检验犯两类错误的概率；
2. 若要使检验犯第二类错误的概率 $\beta \le 0.01, n$ 最小应取多少?
3. 证明当 $n \to \infty, \alpha \to 0, \beta \to 0.$

解答：

1. 由定义可知，犯第一类错误的概率为

$$\alpha = \mathbb{P}(\bar{X} > 2.6|H_0) = \mathbb{P}(\frac{barX - 2}{\sqrt{1/20}} > \frac{2.6-2}{\sqrt{1/20}} = 1 - \Phi(2.68) = 0.0037.$$

这是因为当原假设 H_0 成立时，$\bar{X} \sim N(2, 1/20)$. 同理犯第二类错误的概率为

$$\beta = \mathbb{P}(\bar{X} < 2.6|H_1) = \mathbb{P}(\frac{barX - 3}{\sqrt{1/20}} < \frac{2.6-3}{\sqrt{1/20}} = \Phi(-1.79) = 0.00367.$$

这是因为当原假设 H_1 成立时，$\bar{X} \sim N(3, 1/20)$.

2. 若要使犯第二类错误的概率

$$\beta = \mathbb{P}(\bar{X} < 2.6|H_1) = \mathbb{P}(\frac{barX - 3}{\sqrt{1/n}} < \frac{2.6-3}{\sqrt{1/n}} = \Phi(-0.4/sqrtn) \le 0.01.$$

即 $1 - \Phi(0.4/\sqrt{n}) \le 0.01$ 或 $\Phi(0.4/\sqrt{n}) > 0.99$, 查表得 $0.4/sqrtn > 2.33$, 由此可得 $n > 33.93$. 即 n 最小取 34，才能保证。

3. 在样本量为 n 时，犯第一类错误的概率为

$$\alpha = \mathbb{P}(\bar{X} > 2.6|H_0) = \mathbb{P}(\frac{barX - 2}{\sqrt{1/n}} > \frac{2.6-2}{\sqrt{1/n}} = 1 - \Phi(0.6\sqrt{n}).$$

可知当 $n \to \infty$ 时，$\alpha \to 0$. 同理犯第二类错误的概率为

$$\beta = \mathbb{P}(\bar{X} < 2.6|H_1) = \mathbb{P}(\frac{barX - 3}{\sqrt{1/n}} < \frac{2.6-3}{\sqrt{1/n}} = \Phi(-0.4/sqrtn).$$

可知当 $n \to \infty$ 时，$\beta \to 0$.

练习 8.2 假设样本 $X_1, X_2, \cdots, X_10$ 来自于伯努利总体 $b(1,p)$, 考虑如下的检验问题

$$H_0: p = 0.2;\ vs\ p = 0.4.$$

若检验由拒绝域 $W := \{\bar{X} > 0.5\}$ 确定。求该检验犯两类错误的概率。

思考 8.3 在假设检验问题中，若检验结果为接受原假设，则可能犯哪一类错误? 若检验结果为拒绝原假设，则又可能犯哪一类错误?

解答:

若实验结果为接受原假设，则可能有两种情况：其一是原假设为真，此时检验结果是正确的，未犯错误。其二是原假设不真，此时检验结果就错了，这种错误是接受了不真的原假设，为第二类错误。

若实验结果为拒绝原假设，则也有两种情况：其一是原假设本身不真，此时检验结果是正确的，未犯错误。其二是原假设是真的，此时检验结果就错了，这种错误是拒绝了真的原假设，为第一类错误。

例 8.4 有一批枪弹，出厂时其初速 $v \sim N(950, 100)$, (单位 m/s) 经过较长时间的储存，取出 9 发进行测试，得到样本值如下:

$$914, 920, 910, 934, 953, 945, 912, 924, 940$$

根据以往经验，储存后枪弹初速仍服从正态分布，且标准差保持不变，问是否可认为这批枪弹初速有所降低?（$\alpha = 0.05$）

解答:

这是一个单侧检验问题，总体 $v \sim N(950, 100)$. 待检验的原假设 H_0 和备择假设 H_1 分别为

$$H_0: \mu = 950\ vs\ H_1: \mu < 950.$$

在显著水平为 α 下，拒绝域为 $\{\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} < u_\alpha\}$. 若取 $\alpha = 0.05$, 查表可得 $u_\alpha = -1.65$. 经计算可得 $\bar{x} = 928$, 因此 $\frac{928-950}{10/3} = -6.6$, 即样本落在拒绝域内，故拒绝原假设，可以认为枪弹的初速有所降低。

Remark: 这里注意，上题的检验问题与检验问题

$$H_0: \mu \geq 950\ vs\ H_1: \mu < 950.$$

具有完全相同的拒绝域，这是因为两者拒绝域的形式 $W = \{\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} < c\}$ 是一样的。

例 8.5 已知某铁厂铁水含碳量服从正态分布 $N(4.55, 0.108^2)$, 现测量 9 炉铁水，其平均含碳量为 4.484，如果铁水含碳量的方差没有变化，可否认为现在铁水含碳量均值仍为 4.55?（$\alpha = 0.05$）

解答:

这是一个关于正态总体均值的双侧检验问题。待检验的原假设 H_0 和备择假设 H_1 分别为

$$H_0 : \mu = 4.55 \; vs \; H_1 : \mu \neq 4.55.$$

因为总体方差已知，在显著水平为 α 下，拒绝域为 $\{|\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}| > u_{1-\alpha/2}\}$. 若取 $\alpha = 0.05$, 查表可得 $u_{1-\alpha/2} = 1.96$. 已知 $\bar{x} = 4.484$, 因此 $\frac{4.484-4.55}{0.108/3} = -1.83$, 即样本没有落在拒绝域内，故不能拒绝原假设，可以认为铁水的碳含量均值没有改变。

例 8.6 若需要对某正态总体的均值进行假设检验

$$H_0 : \mu = 15 \; vs \; H_1 : \mu < 15.$$

已知 $\sigma^2 = 2.5$, 取 $\alpha = 0.05$. 若要使 H_1 中当 $\mu \leq 13$ 时，犯第二类错误的概率不超过 0.05，求所需的样本容量。

解答:

由于正态总体的方差已知，对于该单侧检验问题，拒绝域的形式为 $\{\frac{\bar{X}-15}{2.5/\sqrt{n}} < u_{0.05}\}$. 若取 $\alpha = 0.05$, 查表可得 $u_\alpha = -1.65$，即检验的拒绝域为 $\{\frac{\bar{X}-15}{2.5/\sqrt{n}} < -1.65\}$. 于是，当 $\mu \leq 13$ 时，犯第二类错误的概率应满足

$$\beta = \mathbb{P}(\frac{\bar{X}-15}{2.5/\sqrt{n}} > -1.65|\mu \leq 13) \leq 0.05.$$

即

$$\begin{aligned}
\beta &= \mathbb{P}(\frac{\bar{X}-15}{2.5/\sqrt{n}} > -1.65) \\
&= \mathbb{P}(\frac{\bar{X}-\mu}{2.5/\sqrt{n}} > -1.65 + \frac{15-\mu}{2.5/\sqrt{n}}) \\
&= 1 - \Phi(-1.65 + \frac{15-\mu}{2.5/\sqrt{n}}) \\
&\leq 0.05
\end{aligned}$$

由于 β 是 μ 的减函数（分布函数的性质），因此只需 $-1.65 + \frac{15-13}{2.5/\sqrt{n}} > 1,65$ 即可，由此可解得 $n \geq 7$.

例 8.7 假定考生的成绩服从正态分布，在某地的一次数学统考中，随机抽取了 36 位考生，算得平均值为 66.5 分，标准差为 15 分。问在显著水平为 0.05 的条件下，能否认为这次考试全体考生的平均成绩为 70 分？

解答：

这是一个关于正态总体均值的双侧检验问题。待检验的原假设 H_0 和备择假设 H_1 分别为

$$H_0 : \mu = 70 \; vs \; H_1 : \mu \neq 70.$$

由于总体的方差未知，故使用 t 检验法。拒绝域为 $W = \{|\frac{\bar{X}-\mu_0}{S/\sqrt{n}}| > t_{1-\alpha/2}(n-1)\}$. 当显著水平 $\alpha = 0.05$ 时，$t_{0.975}(35) = 2.0301$. 由已知条件 $\bar{x} = 66.5, s = 15$, 因此检验统计量的值为

$$\frac{66.5 - 70}{15/6} = -1.4$$

故接受原假设，认为全体考生的平均值为 70 分。

例 8.8 设在木材中抽取 100 根，测其小头直径，得到样本平均数 $\bar{x} = 11.2$ cm, 样本标准差为 $s = 2.6$ cm. 问能否认为这批木头的直径不低于 12 cm？（取 $\alpha = 0.05$）

解答：

Remark: 这儿没有总体是正态的假设，所以肯定用大样本性质，即中心极限定理。有时即使有正态的假设，t 分布的分位数也只有有数的几个给出了，在样本量比较大，且没有对应的 t 分位数的时候也应使用大样本方法。

由于本题的样本量较大，可以认为样本均值服从正态分布，依题意，需要建立的原假设和备择假设分别为

$$H_0 : \mu \geq 12 \; vs \; \mu < 12.$$

可知拒绝域的形式为 $\{\frac{\sqrt{n}(\bar{X}-12)}{S} < u_\alpha\}$, 若取 $\alpha = 0.05$, 查表可得 $u_\alpha = -1.65$，将样本值带入检验统计量

$$\frac{10(11.2 - 12)}{2.6} = -3.0769 < -1.65.$$

即样本落在拒绝域内，故拒绝原假设，可以认为木材小头的直径均值小于 12 cm.

例 8.9 从某锌矿东西两条支脉中，分别抽取样本容量为 9 和 8 的样本进行测试，得到样本均值和样本方差如下：

东支： $\bar{x}_1 = 0.230, s_1^2 = 0.1337;$

西支： $\bar{x}_2 = 0.269, s_2^2 = 0.1736.$

若东西两支锌矿都服从正态且方差相同，问东西两支矿脉的含锌量的均值是否可以看做一样？（取 $\alpha = 0.05$）

解答：

依题意，可知需要检验的一对假设是

$$H_0 : \mu_1 = \mu_2 \; vs \; \mu_1 \neq \mu_2.$$

这是一个方差未知的两正态总体的双边检验问题，因此拒绝域为 $W = \{|\frac{\bar{X}_1-\bar{X}_2-(\mu_1-\mu_2)}{S_\omega\sqrt{1/n_1+1/n_2}}| > t_{1-\alpha/2}(n_1+n_2-2)\}$，由样本数据可以算得

$$s_w = \sqrt{\frac{8.1337+7.1736}{9+8-2}} = 0.3903.$$

因此检验统计量为

$$\frac{0.230-0.269}{0.3903\sqrt{1/9+1/8}} = -0.2506.$$

当 $\alpha = 0.05$ 时，$t_{0.975}(15) = 2.1314 > 0.2506$，因此接受原假设，即认为两矿脉含锌量的均值相同。

例 8.10 一药厂生产一种新的止痛片，厂方希望验证服用新药片后起作用的时间比原来至少缩减了一半，因此厂方提出需要检验

$$H_0 : \mu_1 \leq 2\mu_2 \; vs \; H_1 : \mu_1 > 2\mu_2.$$

其中，μ_1, μ_2 分别为原止痛片和新生产止痛片产生作用时间的总体的均值。假设两总体均为正态且方差已知分别为 σ_1^2, σ_2^2. 现分别从两个总体中取一个样本 $x_1, x_2, \cdots, x_n$ 和 $y_1, \cdots, y_m$，设两个样本相互独立。试给出上述检验问题的检验统计量和拒绝域。

解答：

设 X 表示原止痛片起作用时间间隔，$x_1, \cdots, x_n$ 为样本。Y 是新生产止痛片起作用时间间隔，$y_1, \cdots, y_m$ 为样本，两个样本相互独立，且 $X \sim N(\mu_1, \sigma_1^2), Y \sim N(\mu_2, \sigma_2^2)$，待检验的一对假设为

$$H_0 : \mu_1 \leq 2\mu_2 \; vs \mu_1 > 2\mu_2.$$

为此，先构造 $\mu_1 - 2\mu_2$ 点估计 $\bar{x} - 2\bar{y}$. 依题意 $\bar{X} - 2\bar{Y} \sim N(\mu_1 - 2\mu_2, \sigma_1^2/n + \sigma_2^2/m)$，由于方差已知，则 $\bar{X} - 2\bar{Y}$ 没有其他未知参数，可以构造检验统计量

$$\frac{\bar{X}-2\bar{Y}}{\sqrt{\sigma_1^2/n+\sigma_2^2/m}}.$$

因此显著水平为 α 的拒绝域为 $W = \{\frac{\bar{X}-2\bar{Y}}{\sqrt{\sigma_1^2/n+\sigma_2^2/m}} > u_{1-\alpha}\}$.

Remark: 思考假如上面例题的方差未知，如何构造检验统计量和拒绝域。

例 8.11 已知维尼纶纤度在正常条件下服从正态分布，且标准差为 0.048. 从某天的产品中抽取 5 根进行检验，测得纤度为

$$1.32, 1.55, 1.36, 1.40, 1.44.$$

问这一天纤度总体的标准差是否合格?（$\alpha = 0.05$）

解答:

这是一个关于正态总体方差的双侧检验问题，待检验的原假设和备择假设分别为

$$H_0 : \sigma^2 = 0.048^2 \ vs \ \sigma^2 \neq 0.048^2.$$

检验统计量为 $\frac{(n-1)S^2}{\sigma^2}$, 此处 $n = 5$, 若取显著水平 $\alpha = 0.05$, 查表得 $\chi^2_{0.025}(4) = 0.4844, \chi^2_{0.975}(4) = 11.1433$. 由样本数据计算可得

$$\frac{(n-1)S^2}{\sigma^2} = \frac{0.031112}{0.048^2} = 13.5609 > 11.1433.$$

因此拒绝原假设，认为这一天标准差不正常。

练习 8.12 某电工器材厂生产一种保险丝，测量其熔断时间，按通常情况方差为 400. 今天从产品中抽取容量为 25 的样本，测得其熔断时间，并计算得 $\bar{x} = 62.24, s^2 = 404.77$, 问这天熔断时间的方差与通常有没有区别?（$\alpha = 0.05$, 假定熔断时间服从正态分布。）

练习 8.13 测得两批电子器件的样品的电阻（单位 Ω）为

$$\begin{aligned} A\text{批} &: 0.140, 0.138, 0.143, 0.142, 0.144, 0.137 \\ B\text{批} &= 0.135, 0.140, 0.142, 0.136, 0.138, 0.140 \end{aligned}$$

设这两批电器的电阻分别服从 $N(\mu_1, \sigma_1^2), N(\mu_2, \sigma_2^2)$, 且两样本独立。

1. 试检验两总体的方差是否相等?（$\alpha = 0.05$）
2. 试检验两总体的均值是否相等?（$\alpha = 0.05$）

例 8.14 某厂用两种不同的原料生产同一类型的产品，随机选取 A 原料生产的产品 22 件，测得其平均质量为 2.36（kg），样本标准差为 0.57（kg）。取使用 B 原料生产的产品 24 件，测得其平均质量为 2.55（kg），样本标准差为 0.48（kg）。设产品质量服

从正态分布，两样本独立，问能否认为用 B 原料生产的产品明显的比原料 A 生产的大。($\alpha = 0.05$)

解答：

设 X 表示原料 A 生产的产品质量，Y 表示原料 B 生产的产品质量，则 $X \sim (\mu_1, \sigma_1^2), Y \sim N(\mu_2, \sigma_2^2)$. 通过问题表述，可知这是关于两正态均值的检验问题，且为了能够显著的认为原料 B 生产的产品比用原料 A 大，需要选择该假设为备择假设，只有拒绝原假设时，才能说明用 B 原料生产的产品明显的比原料 A 生产的大。因此，可以建立如下的假设检验问题

$$H_0 : \mu_1 \geq \mu_2 \; vs \; \mu_1 < \mu_2.$$

为了完成这个检验，应该先对方差是否相等进行检验，如果 $\sigma_1^2 = \sigma_2^2$, 则可以使用 t 统计量作为检验统计量，否则只能用其他方法（近似 t 检验，或者样本量较大时，用中心极限定理）。

对于检验问题

$$H_{00} : \sigma_1^2 = \sigma_2^2 \; vs \; H_{11} : \sigma_1^2 \neq \sigma_2^2.$$

可以使用检验统计量 $F = \frac{s_X^2}{s_Y^2} = \frac{0.57^2}{0.48^2} = 1.4101$. 如果取 $\alpha = 0.05$, 查表得 $F_{0.025}(21, 23) = 0.4201, F_{0.975}(21, 23) = 2.3404$. 观测值未落入拒绝域，因此可以认为方差相等。

在两总体方差相等的假设下，可以使用 t 检验。由样本数据可以算得

$$s_w = \sqrt{\frac{21.57^2 + 23.48^2}{22 + 24 - 2}} = 0.5249.$$

因此检验统计量为

$$\frac{2.55 - 2.36}{0.5249\sqrt{1/22 + 1/24}} = 1.2264.$$

当 $\alpha = 0.05$ 时，$t_{0.975}(44) = 1.645 > 1.2664$, 因此接受原假设，即认为用 B 原料生产的产品没有明显的比原料 A 生产的大.

例 8.15 设总体 $X \sim N(\mu_1, \sigma^2)$, 总体 $Y \sim N(\mu_2, \sigma^2)$, 从总体 X 抽取样本 $x_1, \cdots, x_m$, 从总体 Y 抽取样本 $y_1, \cdots, y_n$, 两样本相互独立。考虑如下检验问题

$$H_0 : c\mu_1 + d\mu_2 = \delta \; vs \; H_1 : c\mu_1 + d\mu_2 \neq \delta.$$

其中 $c \neq 0, d \neq 0, \delta$, 都是已知常数，求检验统计量和拒绝域。

解答：

以 $\bar{x}, \bar{y}$ 表示两个样本的均值，s_x^2, s_y^2 表示两个样本的样本方差，记 $s_\omega = \frac{(m-1)s_x^2 + (n-1)s_y^2}{n+m-2}$. 由所给的条件可知

$$\bar{x} \sim N(\mu_1, \sigma^2/m), \bar{y} \sim N(\mu_2, \sigma^2/n).$$

因此

$$c\bar{x} + d\bar{y} \sim N(c\mu_1 + d\mu_2, (c^2/m + d^2/n)\sigma^2).$$

在原假设成立时，可知

$$\frac{c\bar{x}+d\bar{y}-\delta}{\sigma\sqrt{c^2/m+d^2/n}} \sim N(0,1).$$

此外

$$\frac{(m-1)s_x^2}{\sigma^2} \sim \chi^2(m-1), \frac{(n-1)s_y^2}{\sigma^2} \sim \chi^2(n-1).$$

且两样本独立，可得

$$\frac{(n+m-2)s_\omega^2}{\sigma^2} \sim \chi^2(n+m-2).$$

因此，在原假设成立时，可以构造检验统计量

$$\frac{c\bar{x}+d\bar{y}-\delta}{s_\omega\sqrt{c^2/m+d^2/n}} \sim t(n+m-2).$$

取显著性水平为 α, 可知拒绝域为

$$W=\{|\frac{c\bar{x}+d\bar{y}-\delta}{s_\omega\sqrt{c^2/m+d^2/n}}|>t_{1-\alpha/2}(n+m-2)\}.$$

例 8.16 某大学随机调查了 120 名男同学，发现有 50 人爱看武侠小说。随机调查了 85 名女同学，发现有 23 人喜欢。用大样本方法在 $\alpha=0.05$ 的水平下确认：男女同学在爱看武侠小说上有无明显差异?

解答:

设 X 为 120 男生中喜欢看武侠小说的人数，p_1 为其真实比例；Y 为 85 名女生中喜欢看武侠小说的人数，p_2 为其真实比例，则 $X \sim b(120,p_1), Y \sim b(85,p_2)$，待检验的问题为

$$H_0: p_1=p_2 \; vs \; p_1 \neq p_2.$$

注意此时样本量比较大，故使用大样本法。由中心极限定理可知

$$\bar{X}\dot{\sim}N(p_1,\frac{p_1(1-p_1)}{120}), \bar{y}\dot{\sim}N(p_2,\frac{p_2(1-p_2)}{85}).$$

在原假设 H_0 成立的条件下，近似有

$$\frac{\bar{x}-\bar{y}}{\sqrt{\frac{\hat{p}(1-\hat{p})}{120}+\frac{\hat{p}(1-\hat{p})}{85}}}\dot{\sim}N(0,1)$$

这里 $\hat{p}=\frac{50+23}{120+85}=0.3561$. 将 $\bar{x},\bar{y}$ 的值带入，可得

$$\frac{0.4167-0.2706}{0.3561(1-0.3561)\sqrt{1/120+1/85}}=2.1522>1.96.$$

故在显著性水平 $\alpha=0.05$, 样本观测值落在了拒绝域，因此拒绝原假设，认为男女生在看武侠小说上有显著差别。

练习 8.17 某人声称他能根据股票价格的历史图标预测未来的涨跌，若在一场测试中，他共做了 10 次预测，报对了 8 次。在显著性水平 0.05 下，能否相信他有这种能力？

例 8.16 某厂生产的元件平均寿命为 1200 h，偏低，现在工厂进行技术革新，革新后任选 8 个元件进行检验，得到样本值

$$2686, 2001, 2082, 792, 1660, 4105, 1416, 2089.$$

假设元件寿命服从指数分布，在显著水平 $\alpha = 0.05$ 下，问革新后元件寿命有无明显提高？

解答：

假设 X 表示元件的寿命，可知 $X \sim Exp(1/\theta)$, 其中 θ 是元件的平均寿命。依题意，可以原假设和备择假设分别是

$$H_0 : \theta \leq 1200 \; vs \; H_1 : \theta > 1200.$$

取检验统计量为 $\frac{2n\bar{X}}{\theta_0}$, 可知拒绝域形如 $W = \{\frac{2n\bar{X}}{\theta_0} > \chi^2_{0.95}(2n)\}$. 具体到本问题，$n = 8$ 以及查表可得 $\chi^2_{0.95}(16) = 26.2962$. 将样本值带入可得

$$\frac{16.875}{1200} = 28.0517 > 26.2962.$$

即样本落入拒绝域，因此可以认为革新后元件寿命有明显提高。