全国重点名校系列

新版

全国硕士研究生招生考试

考研专业课精品资料

【电子书】2024年北京大学

432统计学考研精品资料

策划：辅导资料编写组

真题汇编 直击考点

考研笔记 突破难点

核心题库 强化训练

模拟试题 查漏补缺

高分学长学姐推荐

【初试】2024年北京大学432统计学考研精品资料

说明：本套资料由高分研究生潜心整理编写，高清PDF电子版支持打印，考研首选资料。

一、北京大学432统计学考研真题汇编

1. **北京大学432统计学[专业硕士]2000-2004年考研真题，暂无答案。**

说明：分析历年考研真题可以把握出题脉络，了解考题难度、风格，侧重点等，为考研复习指明方向。

二、2024年北京大学432统计学考研资料

2. **《概率论与数理统计教程》考研相关资料**

（1）《概率论与数理统计教程》[笔记+课件+提纲]

①北京大学432统计学之《概率论与数理统计教程》考研复习笔记。

说明：本书重点复习笔记，条理清晰，重难点突出，提高复习效率，基础强化阶段首选资料。

②北京大学432统计学之《概率论与数理统计教程》本科生课件。

说明：参考书配套授课PPT课件，条理清晰，内容详尽，版权归属制作教师，本项免费赠送。

③北京大学432统计学之《概率论与数理统计教程》复习提纲。

说明：该科目复习重难点提纲，提炼出重难点，有的放矢，提高复习针对性。

（2）《概率论与数理统计教程》考研核心题库（含答案）

①北京大学432统计学考研核心题库之解答题精编。

②北京大学432统计学考研核心题库之计算题精编。

说明：本题库涵盖了该考研科目常考题型及重点题型，根据历年考研大纲要求，结合考研真题进行的分类汇编并给出了详细答案，针对性强，是考研复习首选资料。

（3）《概率论与数理统计教程》考研模拟题[仿真+强化+冲刺]

①2024年北京大学432统计学考研专业课五套仿真模拟题。

说明：严格按照本科目最新专业课真题题型和难度出题，共五套全仿真模拟试题含答案解析。

②2024年北京大学432统计学考研强化五套模拟题及详细答案解析。

说明：专业课强化检测使用。共五套强化模拟题，均含有详细答案解析，考研强化复习首选。

③2024年北京大学432统计学考研冲刺五套模拟题及详细答案解析。

说明：专业课冲刺检测使用。共五套冲刺预测试题，均有详细答案解析，最后冲刺首选资料。

三、电子版资料全国统一零售价

3. **本套考研资料包含以上一、二部分（高清PDF电子版，不含教材），全国统一零售价：[￥]**

特别说明：

①本套资料由本机构编写组按照考试大纲、真题、指定参考书等公开信息整理收集编写，仅供考研复习参考，与目标学校及研究生院官方无关，如有侵权、请联系我们将立即处理。

②资料中若有真题及课件为免费赠送，仅供参考，版权归属学校及制作老师，在此对版权所有者表示感谢，如有异议及不妥，请联系我们，我们将无条件立即处理！

四、2024年研究生入学考试指定/推荐参考书目（资料不包括教材）

4．**北京大学432统计学考研初试参考书**

茆诗松《概率论与数理统计教程》，2011年

五、本套考研资料适用学院和专业

数学科学学院

版权声明

目录

2024年北京大学432统计学备考信息

北京大学432统计学考研初试参考书目

茆诗松《概率论与数理统计教程》，2011年

北京大学432统计学考研招生适用院系

数学科学学院

北京大学 432 统计学历年真题汇编

北京大学 432 统计学 2000 年考研真题（暂无答案）

北京大学 2000 年研究生入学考试统计学试题

1、 设简单随机样本(XI，Y1)， (X2，Y2)，……(Xn，Yn)满足等式

$$\sum_{i=1}^{n} X_i = \sum_{i=1}^{n} Y_i = 0, \sum_{i=1}^{n} X_i^2 = \sum_{i=1}^{n} Y_i^2 = 0$$

而且 X 和 Y 之间的样本相关系数为 r，试写出 Y 对 X 的回归方程。 (10 分)

2、 某市场研究机构从一城市居民中随机选择了 2000 户邮寄调查问卷，收回问卷 890 份，按照家庭收入分组情况如下表所示。另外通过已有资料已知该城市中各种收入范围的居民所占百分比现在该市场研究机构需要知道这些问卷是否只有代表性。请使用一种统计检验方法并给出结论。 (取显著水平为 0．05) (如有必要，可以参照试卷最后所附的统计表) (15 分)

家庭收入（元）	回收问卷的家庭次数分布	该收入家庭数占城市总家庭数百分比
<500	121	12
500~1000	166	20
1000~2000	313	38
2000~5000	197	23
>=5000	93	7
总计	890	100

3、 某市场的年度销售额数据如下表所示。

年份	销售额（万元）	一次移动平均	二次移动平均
1990	8.60		
1991	10.00		
1992	14.80	11.13	
1993	13.70	12.83	
1994	16.20	14.90	12.96
1995	19.30	16.40	14.71
1996	22.80	19.43	16.91
1997	23.10	21.73	19.19
1998	25.90	23.93	21.70
1999	29.00	26.00	23.89
2000			

（注：图形被省略）

表中：同时列出了 3 年的一次、二次移动平均数。

(1)根据表中提供的资料，试求 2000 年利 2001 年该商场销售额的预测值。

(2)如果使用一次指数平滑方法，选择平滑系数等于多少将导致与上面的一次移动平均值具有相同的滞后效应。

(3)假设我们得到的原始资料是每年的各个季度的销售额，而且根据已有数据初步算得每个季度的季节指数分别为 95%，123%，110%，82%。请利用①的结果预测 2000 年各个季度的销售额。 (15 分)

4、 设 X1，X2，…Xn 是来自在区间[0， θ]上面均匀分布的总体的简单随机样本(θ>0)。试求 θ 的最大似然估量 θ(加尖)，并判别 θ(加尖)是否是 θ 的无偏估计。 (10 分)

5、 兹有一家房地产有限公司委托某咨询公司对北京市的高档商品房需求状况进行抽样调查，试问：

a. 抽样的类型及其应用的范围(简述之)。(20 分)

b. 如果采用简单随机抽样，当置信度等于 0．95 时，如按不重复抽样方法抽取的必要样本单位数是多少? (按历史成数方差最大值计算：其允许：误差范围为±4．9%) (10 分)

6、 某管理局所属两个工厂在某年第一、二季度的产值和工人人数资料如下(20 分)

厂名	产值(万元) 第一季度	第二季度	工人数(人) 第一季度	第二季度
一厂	60	88	600	800
二厂	24	12	400	200

试利用统计指数法分析全管理局劳动生产率变化的原因。

附表：常用分布的 α 上侧分位数 X_α (即 $P(X \ge X_\alpha) = \alpha$ 表。 (此表略)

北京大学 432 统计学 2001 年考研真题（暂无答案）

2001 年硕士研究生入学考试试题

一、（45 分）简要回答下列问题：

1、（5 分）统计指标的主要特点是什么？

2、（10 分）在统计中，测定标志变异性的方法有极差、平均差、标准差和标志变动系数等几种主要方法，请用计算公式表示，并分别说明它们的作用和局限性。

3、（10 分）在计算价格指数时，“拉斯贝尔价格指数”和“派许价格指数”有什么不同（用计算公式表示）？各反映什么问题？你认为哪一个公式更具有现实经济意义，为什么？

4、（5 分）增加值和净产值；社会总产值和国内生产总值有什么区别，请具体说明。

5、（5 分）请说明经济效益的基本含义。

6、（10 分）在对单个总体均值的假设 $H_0: u=u_0$ 作 t 检验时，计算机软件得到的 P 值是 0.074，请你解释此时 P 值的含义。如果计算发现样本均值 $\bar{x}>u_0$，现在需要进一步判断是否有充分的理由认为总体均值是否也大于 u_0，你该如何去做？根据刚才的 P 值，你对“总体均值是否也大于” u_0 可以获得什么结论？

二、（15 分）计算下列问题

1、（10 分）某电灯泡厂对 2000 个产品进行使用寿命检验，随机抽取 2%进行测试，如果电灯泡平均使用时间为 1057.63 小时（样本指标 $\bar{x}$），电灯泡合格率为 91.75%（样本成数），重复抽样电灯泡平均使用时间抽样平均误差为 2.6535 小时，重复抽样电灯泡合格率抽样平均误差为 1.376%，在 68.27%和在 95.45%的概率保证下，求平均数和成数的抽样极限误差。

2、（5 分）在整理我国 1982 年第三次人口普查资料时采用了如下的异距分组：

不满周岁	1-3 岁	4-6 岁	7-12 岁	13-15 岁
16-17 岁	18-19 岁	20-24 岁	25-29 岁	30-40 岁
35-39 岁	40-44 岁	45-49 岁	50-54 岁	55-59 岁
60-64 岁	65-79 岁	80-99 岁	100 岁及以上	

试问：1）为什么要采用异距数列，为的是说明什么问题？

2）第一组的下限和最后一组的上限请分别标明。

三、（15 分）设 X1，X2，… Xn 是独立同分布的，具有密度函数 $f(x;u,\delta)=\frac{1}{\delta}\exp(-\frac{x-u}{\delta}), x\ge u, -\infty<u<+\infty, \delta>0$ 试求 u 和 δ 的最大似然估计。

四、（15 分）一项民意测验表明，通过抽取简单随机样本调查的 840 名男子和 750 名女子中分别有 16%和 25%赞成立法在全国范围内禁止销售烈酒。

1、请你写出男女中赞成该项立法的比例之差的 95%的置信区间；

2、有报道说“男子赞成这项法案的比例比女性低 15%”，你根据这项民意测验的结果能否接受该报道的说法，陈述你的理由。

五、（10 分）对于二元线性回归模型 $y=\beta_0+\beta_1x_1+\beta_2x_2+\varepsilon$，假定利用样本（$x_{i1},x_{i2};y_i$）$i=1,2,\cdots,n$ 通过最小二乘估计得到的回归方程是 $\hat{y}=b_0+b_1x_1+b_2x_2$，判决系数是 R^2，记 $\hat{y}=b_0+b_1x_{i1}+b_2x_{i2}$，$i=1,2,\cdots,n$，试证（$y_i,\hat{y}_i$），$i=1,2,\cdots,n$ 的样本相关系数的平方是 R^2。

启用前机密 北京大学 2001 年硕士研究生入学考试试题

考试科目：统计学　　　　　考试时间：1 月 14 日下午

招生专业：统计学　　　　　研究方向：管理统计

试　　题：（注意:答案一律写在答题纸，否则不计分）

一、（45 分）简要回答下列问题：

1、（5 分）统计指标的主要特点是什么？

2、（10 分）在统计中，测定标志变异性的方法有极差、平均差、标准差和标志变动系数等几种主要方法，请用计算公式表示，并分别说明它们的作用和局限性。

3、（10 分）在计算价格指数时，“拉斯贝尔价格指数”和“派许价格指数”有什么不同（用计算公式表示）？各反映什么问题？你认为哪一个公式更具有现实经济意义，为什么？

4、（5 分）增加值和净产值；社会总产值和国内生产总值有什么区别，请具体说明。

5、（5 分）请说明经济效益的基本含义。

6、（10 分）在对单个总体均值的假设 $H_0: \mu = \mu_0$ 作 t 检验时，计算机软件得到的 P 值是 0.074，请你解释此时 P 值的含义。如果计算发现样本均值 $\bar{x} > \mu_0$，现在需要进一步判断是否有充分的理由认为总体均值也大于 μ_0，你该如何去做？根据刚才的 P 值，你对“总体均值是否也大于 μ_0 可以获得什么结论？

二、（15 分）计算下列问题

1、（10 分）某电灯泡厂对 2000 个产品进行使用寿命检验，随机抽取 2%进行测试，如果电灯泡平均使用时间为 1057.63 小时（样本指标 $\bar{x}$），电灯泡合格率为 91.75%（样本成数），重复抽样电灯泡平均使用时间抽样平均误差为 2.6535 小时，重复抽样电灯泡合格率抽样平均误差为 1.376%，在 68.27%和在 95.45%的概率保证下，求平均数和成数的抽样极限误差。

2、（5 分）在整理我国 1982 年第三次人口普查资料时采用了如下的异距分组：

不满周岁　1-3 岁　4-6 岁　7-12 岁　13-15 岁

2024 年北京大学 432 统计学考研核心笔记

《概率论与数理统计教程》考研核心笔记

第 1 章　随机事件与概率

考研提纲及考试要求

考点：随机现象
考点：样本空间
考点：随机事件
考点：随机变量
考点：事件间的关系

考研核心笔记

【核心笔记】随机事件及其运算

1. 随机现象

概率论与数理统计的研究对象是随机现象.
定义在一定条件下，并不总是出现相同结果的现象称为随机现象.
随机现象的特点：
结果不止一个；
哪一个结果出现事先不知道.
随机试验：在相同条件下可以重复的随机现象.
概率论与数理统计主要研究能大量重复的随机现象.

2. 样本空间

定义：随机现象的一切可能结果组成的集合称为样本空间，记为$\Omega=\{\omega\}$
其中，ω表示基本结果，称为样本点.
离散样本空间和连续样本空间：
样本点个数为有限或可列的样本空间称为离散样本空间. 如Ω_1，Ω_2，Ω_3.
样本点个数为不可列无限个的样本空间称为连续样本空间. 如Ω_4，Ω_5.

3. 随机事件

定义随机现象的某些样本点组成的集合称为随机事件，简称事件.
事件的图形表示称为维恩图.
由样本空间中的单个元素组成的子集称为基本事件.
Ω称为必然事件，$\varnothing$称为不可能事件.

4. 随机变量

定义用来表示随机现象结果的变量称为随机变量.

随机变量常用大写字母 X, Y, Z 来表示，很多随机事件可用随机变量表示.

5. 事件间的关系

事件之间的关系包括包含关系、相等关系、互不相容关系等，以及各种关系的维恩图表示.

包含关系

若 $A \subset B$，则称事件 B 包含事件 A，即事件 A 发生必导致事件 B 发生.

二相等关系

若 $A \subset B$，且 $B \subset A$，即 $A = B$，称事件 A 与事件 B 相等.

三互不相容

若 $A \cap B = \Phi$，则称事件 A 与事件 B 是互不相容的，或互斥的.即事件 A 与事件 B 不能同时发生.

6. 事件运算

（1）事件运算：并、交、差、余.

①事件 A 与 B 的并记为 $A \cup B$，$A \cup B = \{x | x \in A 或 x \in B\}$，

当且仅当 A 与 B 至少有一个发生时，事件 $A \cup B$ 发生.

②事件 A 与 B 的交记为 $A \cap B$ 或 AB $A \cap B = \{x | x \in A 且 x \in B\}$，当且仅当 A 与 B 同时发生时，事件 $A \cap B$ 发生.

事件的并交运算可推广到有限个或可列个事件，$\bigcup_{k=1}^{n} A_k$ 称为有限并，$\bigcup_{k=1}^{\infty} A_k$ 称为可列并，$\bigcap_{k=1}^{n} A_k$ 称为有限交，$\bigcap_{k=1}^{\infty} A_k$ 称为可列交.

③事件 A 对 B 的差记为 $A - B$，$A - B = \{x | x \in A 且 x \notin B\}$，当且仅当 A 发生、B 不发生时，事件 $A - B$ 发生.

④对立事件 A 的对立事件记为 $\overline{A}$，即 $\overline{A} = \Omega - A$.

事件 A 与 B 互为对立事件的充要条件为若 $A \cup B = \Omega$ 且 $A \cap B = \Phi$

利用对立事件，可知 $A - B = A\overline{B}$

⑤事件的运算性质

a. 交换律：

$$A \cup B = B \cup A, A \cap B = B \cap A;$$

b. 结合律：

$$(A \cup B) \cup C = A \cup (B \cup C), (A \cap B) \cap C = A \cap (B \cap C);$$

c. 分配律：

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C), (A \cap B) \cup C = (A \cup C) \cap (B \cup C);$$

b. 对偶律（德莫根公式）：

$$\overline{A \cup B} = \overline{A} \cap \overline{B}, \overline{A \cap B} = \overline{A} \cup \overline{B}.$$

7. 事件域

（1）定义设 Ω 为一样本空间，$\mathbb{F}$ 为 Ω 的某些子集组成的集合，如果 $\mathbb{F}$ 满足：

① $\Omega \in \mathbb{F}$;

②若 $A \in \mathbb{F}$，则 $\overline{A} \in \mathbb{F}$；

③若 $A_n \in \mathbb{F}$， $n=1,2,\cdots$, 则 $\bigcup_{n=1}^{+\infty} A_n \in \mathbb{F}$.则称 $\mathbb{F}$ 为一事件域或 $\sigma-$ 代数.

（2）常见事件域

（3）波雷尔事件域：$\mathbb{F}=\sigma((-\infty,x),x\in\Re)$.

【核心笔记】概率的定义及其确定方法

1. 概率的公理化定义

定义设 Ω 为一样本空间, $\mathbb{F}$ 为 Ω 上的某些子集组成的一个事件域，如果对任意事件, $A\in\mathbb{F}$ 定义在 $\mathbb{F}$ 上的一个实值函数 P（A）满足：

非负性公理：$P(A)\geq 0;$

正则性公理：$P(A)=1;$

可列可加性公理：若 $A_1,A_2,\cdots A_n,\cdots$ 两两互不相容，有

$$P(\bigcup_{n=1}^{+\infty} A_n)=\sum_{n=1}^{+\infty} P(A_n);$$

则称 P（A）为事件 A 的概率，称三元素 $(\Omega,\mathbb{F},P)$ 为概率空间.

概率是关于事件的满足如上三条公理的函数.

2. 排列与组合公式

两大计数原理

（1）乘法原理完成一件工作分 m 个步骤，第一步骤有 n_1 种方法，第二步骤有 n_2 种方法，……第 m 个步骤有 n_m 种方法，那么完成这件工作共有 $n_1\times n_2\times\cdots\times n_m$ 种方法.

（2）加法原理：完成一件工作有 m 个独立的途径，第一个途径有 n_1 种方法，……第 m 个途径有 n_m 种方法，那么完成这件工作共有 $n_1+n_2+\cdots+n_m$ 种方法.

排列组合公式

以上述两个原理为基础，可以推导出如下的排列、组合等公式：

①排列：从 n 个元素中取出 r 个来排列，既要考虑每次取到哪个元素，又要考虑取出的顺序，这种排列称为选排列，总数 $A_n^r=n(n-1)\cdots\cdots(n-r+1)$，特别地 $A_n^n=n(n-1)\cdots 3\cdot 2\cdot 1=n!$ 称为 n 个元素的全排列.

②重复排列每次选取都是在全体元素中进行，同一元素可被重复选中，这种排列称为有重复排列，总数为 n^r 种.

③组合

从 n 个元素中取出 r 个元素的组合是不考虑元素的顺序的，其组合总数为

$$\binom{n}{r}=\frac{A_n^r}{r!}=\frac{n!}{r!(n-r)!}$$

重复组合

从 n 个元素中每次取一个，放回后再取下一个，连续取 r 次所的组合数为 C_{n+r-1}^r ，称为重复组合数.

3. 确定概率的频率方法

定义在 n 次独立重复试验中，记 $n(A)$ 为事件 A 出现的次数，又称 $n(A)$ 为事件 A 的频数，称

$$f_n(A)=\frac{n(A)}{n}$$

为事件 A 出现的频率.

基本思想在与考察事件 A 有关的随机现象可大量重复进行的条件下，记事件 A 的频率为 $f_n(A)$，随着 n 的增加，$f_n(A)$ 会稳定在一常数 α 附近，这个频率的稳定值就是所求事件 A 的概率.

4. 确定概率的古典方法

基本思想

（1）所涉及的随机现象只有有限个样本点，譬如 n 个；

（2）每个样本点发生的可能性相等；

（3）若事件 A 含有 k 个样本点，则 A 的概率为

$$P(A)=\frac{\text{事件A所含样本点个数}}{\Omega\text{中所含样本点个数}}=\frac{\mathrm{k}}{\mathrm{n}}.$$

5. 确定概率的几何方法

基本思想

（1）如果一个随机现象的样本空间 Ω 充满某个区间，其度量可用 S_Ω 表示；

（2）任意一点落在度量相同的子区域内是等可能的；

（3）若事件 A 为 Ω 中的某个子区域，其度量为 S_A ，则事件 A 的概率为

这时实际向大家介绍了一个很有用的计算方法，即若我们想要计算一个感兴趣的量（上面这个量是 π），则可适当地设计一个随机试验，使试验下某个事件的概率与感兴趣的那个有关，然后重复试验多次，以频率代事件的概率便可求出那个量的近似解来。人们称这种计算方法为随机模拟法或蒙特——卡洛（Monto—Carlo）方法.

6. 确定概率的主观方法

定义统计界的贝叶斯学派认为：一个事件的概率是人们根据经验对该事件发生的可能性给出的个人信念。这样给出的概率称为主观概率。

【核心笔记】概率的性质

1. 概率的可加性

2024年北京大学432统计学考研辅导课件

《概率论与数理统计教程》考研辅导课件

第一章　随机事件与概率

第一章　随机事件与概率

§1.2　概率的定义及其确定方法

- 直观定义 —— 事件A出现的可能性大小.
- 统计定义 —— 事件A在大量重复试验下出现的频率的稳定值称为该事件的概率.
- 古典定义；几何定义.

1.2.1　概率的公理化定义

- 非负性公理：　$P(A)\geq 0$;
- 正则性公理：　$P(\Omega)=1$;
- 可列可加性公理：若$A_1, A_2, \ldots\ldots, A_n \ldots\ldots$

互不相容，则 $P\left(\bigcup_{i=1}^{\infty} A_i\right)=\sum_{i=1}^{\infty} P(A_i)$

1.2.3　确定概率的频率方法

- 随机试验可大量重复进行.
- 进行n次重复试验，记 $n(A)$ 为事件A的频数，
  称 $f_n(A)=\frac{n(A)}{n}$ 为事件A的频率.
- 频率$f_n(A)$会稳定于某一常数(稳定值).
- 用频率的稳定值作为该事件的概率.

1.2.4　确定概率的古典方法

古典方法　设Ω为样本空间，若

① Ω只含有限个样本点;

② 每个样本点出现的可能性相等，

则事件A的概率为:

$P(A)=A$中样本点的个数 / 样本点总数

1.2.5　确定概率的几何方法

若 ① 样本空间Ω充满某个区域，其度量(长度、面积、体积)为S_Ω;

② 落在Ω中的任一子区域A的概率，只与子区域的度量S_A有关，而与子区域的位置无关(等可能的).

则事件A的概率为: $P(A)=S_A/S_\Omega$

几何方法的例子

例1.2.3　蒲丰投针问题

平面上画有间隔为d的等距平行线，向平面任意投掷一枚长为l的针，求针与平行线相交的概率.

蒲丰投针问题(续1)

解：以x表示针的中点与最近一条平行线的距离，又以φ表示针与此直线间的交角.

易知样本空间Ω满足:

$0\leq x\leq d/2$;　$0\leq\varphi\leq\pi$.

Ω形成x-φ平面上的一个矩形，其面积为:

$S_\Omega=d(\pi/2)$.

蒲丰投针问题(续2)

A =“针与平行线相交” 的充要条件是:

$x \le l/2 \sin \varphi$.

针是任意投掷的，所以这个问题可用几何方法求解得

$$P(A)=\frac{S_A}{S_\Omega}=\frac{\int_0^\pi \frac{l}{2}\sin\varphi d\varphi}{d(\pi/2)}=\frac{2l}{d\pi}$$

§1.3 概率的性质

性质1.3.1 $P(\varphi)=0$.

注意: 逆不一定成立.

1.3.1 概率的可加性

性质1.3.2 (有限可加性)

若$AB=\varphi$，则

$P(A\cup B)=P(A)+P(B)$.

可推广到 n 个互不相容事件.

性质1.3.3 (对立事件公式)

$P(\bar{A})=1-P(A)$.

彩票问题——幸运35选7

➢ 购买：从01，……，35 中选7个号码.

➢ 开奖：7个基本号码，1个特殊号码.

中奖规则

1) 7个基本号码
2) 6个基本号码 ＋ 1个特殊号码
3) 6个基本号码
4) 5个基本号码 ＋ 1个特殊号码
5) 5个基本号码
6) 4个基本号码 ＋ 1个特殊号码
7) 4个基本号码，或 3个基本号码 ＋ 1个特殊号码

中奖概率

➢ Ω 中所含样本点个数：C_{35}^7

➢ 将35个号分成三类：

7个基本号码、1个特殊号码、27个无用号码

➢ 记 p_i 为中 i 等奖的概率。利用抽样模型得：

$$p_1=\frac{C_7^7C_1^0C_{27}^0}{C_{35}^7}、\quad p_2=\frac{C_7^6C_1^1C_{27}^0}{C_{35}^7}$$

➢ 中奖概率如下：

$$p_1=\frac{1}{6724520},\quad p_2=\frac{7}{6724520},\quad p_3=\frac{189}{6724520}$$

$$p_4=\frac{567}{6724520},\quad p_5=\frac{7371}{6724520},\quad p_6=\frac{12285}{6724520}$$

$$p_7=\frac{204750}{6724520},$$

➢ 不中奖的概率为：

$$p_0=1-p_1-p_2-p_3-p_4-p_5-p_6-p_7 \quad \frac{6499350}{6724520} \quad 0.966515.$$

生日问题

➢ 求n个人（n小于等于365）中至少有两人生日相同的概率.

➢ 看成 n 个球放入 N=365个盒子中.

➢ P(至少两人生日相同)=1−P(生日全不相同)

➢ 用盒子模型得：$p_n=P$(至少两人生日相同)=

$$1-\frac{365!}{365^n(365-n)!}$$

p_{20}=0.4058，p_{30}=0.6963，p_{50}=0.9651，p_{60}=0.9922

1.4.1 条件概率的定义

定义1.4.1

对于事件A、B，若$P(B)>0$，则称

$P(A|B)=P(AB)/P(B)$

为在B出现的条件下，A出现的条件概率.

条件概率 $P(A|B)$ 的计算

1) 缩减样本空间：将Ω缩减为$\Omega_B=B$.

2) 用定义：$P(A|B)=P(AB)/P(B)$.

例1.4.1 10个产品中有7个正品、3个次品，从中不放回地抽取两个，已知第一个取到次品，求第二个又取到次品的概率.

解：设A={第一个取到次品}，

B={第二个取到次品}，

$P(B|A)=P(AB)/P(A)=(1/15)/(3/10)=2/9$

条件概率的三大公式

- 乘法公式;
- 全概率公式;
- 贝叶斯公式.

1.4.2 乘法公式

性质1.4.2

(1) 若$P(B)>0$，则 $P(AB)=P(B)P(A|B)$;

若$P(A)>0$，则 $P(AB)=P(A)P(B|A)$.

(2) 若$P(A_1A_2\cdots\cdots A_{n-1})>0$，则

$P(A_1A_2\cdots\cdots A_n)$

$=P(A_1)P(A_2|A_1)\cdots\cdots P(A_n|A_1A_2\cdots\cdots A_{n-1})$

1.4.3 全概率公式

性质1.4.3

若事件$B_1, B_2, \cdots\cdots, B_n$是样本空间$\Omega$的一组分割，且$P(B_i)>0$，则

$$P(A)=\sum_{i=1}^{n}P(AB_i)=\sum_{i=1}^{n}P(B_i)P(A|B_i)$$

注意点(1)

- 全概率公式用于求复杂事件的概率.
- 使用全概率公式关键在于寻找另一组事件来"分割"样本空间.
- 全概率公式最简单的形式：

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

例1.4.2 设10件产品中有3件不合格品，从中不放回地取两次，每次一件，求取出的第二件为不合格品的概率。

解：设A="第一次取得不合格品"，

B="第二次取得不合格品".

由全概率公式得：

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

$=(3/10)\times(2/9)+(7/10)\times(3/9)=3/10$

1.4.4 贝叶斯公式

- 乘法公式是求“几个事件同时发生”的概率；
- 全概率公式是求“最后结果”的概率；
- 贝叶斯公式是已知“最后结果”，求“原因”的概率.

已知“结果”，求“原因”

某人从甲地到乙地，乘飞机、火车、汽车迟到的概率分别为0.1、0.2、0.3，他等可能地选择这三种交通工具。若已知他最后迟到了，求他分别是乘飞机、火车、汽车的概率.

（1/6， 2/6， 3/6）

贝叶斯（Bayes）公式

若事件$B_1, B_2, \cdots\cdots, B_n$是样本空间Ω的一组分割，且$P(A)>0$, $P(B_i)>0$，则

$$P(B_i|A)=\frac{P(AB_i)}{P(A)}=\frac{P(B_i)P(A|B_i)}{P(A)}$$

$$=\frac{P(B_i)P(A|B_i)}{\sum_{j=1}^{n}P(B_j)P(A|B_j)} \qquad i=1,2,\ldots,n$$

注 意 点

1) $B_1, B_2, \ldots, B_n$可以看作是导致A发生的原因;
2) $P(B_j|A)$是在事件A发生的条件下, 某个原因B_j发生的概率, 称为“后验概率”;
3) Bayes公式又称为“后验概率公式”或“逆概公式”;
4) 称$P(B_j)$为“先验概率”.

§1.5 独立性

事件的独立性

直观说法：对于两事件，若其中任何一个事件的发生不影响另一个事件的发生，则这两事件是独立的.

$\Leftrightarrow$ $P(A|B) = P(A)$

$\Leftrightarrow$ $P(AB)/P(B) = P(A)$

$\Leftrightarrow$ $P(AB) = P(A)P(B)$

1.5.1 两个事件的独立性

- 定义1.5.1 若事件 A 与 B 满足：
  $P(AB)=P(A)P(B)$，
  则称A与B相互独立，简称A与B独立.
- 结论 A、B 为两个事件，若 $P(A)>0$，则 A 与 B 独立等价于 $P(B|A)=P(B)$.
- 性质1.5.1 若事件A与B独立，则
  A 与 独立、 与 B独立、 与 独立.

1.5.2 多个事件的相互独立性

- 对于A、B、C三个事件，称满足：
  $P(AB)=P(A)P(B)$, $P(AC)=P(A)P(C)$, $P(BC)=P(B)P(C)$
  为A、B、C 两两独立.
- 称满足：$P(ABC)=P(A)P(B)P(C)$
  为A、B、C三三独立.

定义1.5.3 若事件 $A_1, A_2, \ldots\ldots, A_n$满足：
两两独立、三三独立、……、n n 独立
则称$A_1, A_2, \ldots\ldots, A_n$相互独立.

一 些 结 论

若A、B、C 相互独立，则

$A\cup B$ 与 C 独立,

$A\cap B$ 与 C 独立,

$A-B$ 与 C 独立.

2024年北京大学432统计学考研复习提纲

《概率论与数理统计教程》考研复习提纲

概率论与数理统计复习提纲

第一章 随机事件与概率

【知识点提示】熟练掌握随机试验，样本空间，样本点，事件与事件的运算，概率的定义与 性质，古典概型，条件概率与乘法原理，事件的独立性，基本知识点如下：

1、 样本空间的概念，随机事件的概念，事件的关系与运算。

2、事件频率的概念，了解概率的统计定义。

3、 概率的古典定义，计算简单的古典概率。

4、 概率的公理化定义，概率的基本性质及概率加法定理。

5、 条件概率的概念、概率的乘法定理、全概率公式和贝叶斯公式。

6、事件的独立性概念。

【重难点提示】事件的运算、概率的定义及性质，事件的独立性。利用概率的性质解决古典 概型的概率及对"相互独立事件"的理解。

第一节. 随机事件及其运算

一、统计规律性

二、随机现象

三、随机试验

四、样本空间

五、随机事件

六、随机事件之间的运算

第二节　概率的定义及其确定方法

一、频率

二、概率的公理化定义和性质

三、确定方法

第三节　概率的性质

一、古典概型的定义

二、古典概型的典型举例和应用

三、性质

第四节　条件概率

一、条件概率

二、乘法公式

三、全概率公式

四、贝叶斯公式

第五节　独立性

一、事件的独立性的定义

二、事件的独立性的应用

第二章 随机变量及其分布

【知识点提示】初步了解随机变量，分布函数及分布函数的性质，离散型随机变量及其概率 分布，连续型随机变量及概率密度函数，随机变量函数的分布，基本知识点如下：

1、 随机变量的概念、离散型随机变量及分布律的概念和性质。

2、 分布函数的概念和性质，利用概率分布计算有关事件的概率。

3、 简单随机变量函数的概率分布。

4、 0-1 分布、二项分布、普哇松分布的定义，知道二项分布与普哇松分布的关系。

5、 均匀分布、指数分布、正态分布与标准正态分布的定义与关系，正态分布的概率密度函数的性质，用正态分布的概率密度函数计算概率问题。

6、 随机变量函数的分布的定理计算方法。

【重难点提示】随机变量与分布函数的概念，离散型随机变量、连续型随机变量的概念及概 率的求法，对分布函数的理解及用该函数求具体概率问题，随机变量函数的分布。

第一节　　随机变量及其分布

一、离散型随机变量

二、连续型随机变量

第二节　　随机变量的数学期望

一、 定义

二、 性质

三、 期望的应用

第三节　　随机变量的方差与标准差

一、 随机变量的方差

二、 随机变量的标准差

第四节　　常用的离散分布

掌握离散分布函数的定理求法

第五节　　常用的连续分布

掌握连续分布的类型及应用

第六节 常用变量函数的分布

单值连续函数的分布特点及应用

第七节 分布的其他特征数

一、定义；

二、性质；

三、应用

第三章 多维随机变量及其分布

【知识点提示】掌握二维随机变量，联合分布，边缘分布，条件分布 ，相互独立的随机变 量，两个随机变量的函数的分布。基本知识点如下；

1、二维随机变量的概念（离散型随机变量及连续型随机变量）及概率密度的概念和性质。

2、 二维分布函数的概念和性质，利用概率分布计算有关事件的概率。

3、 二元随机变量的概念，联合分布与边缘分布的概念及其关系，离散型和连续型二维 随机变量的条件分布的计算。

4、 随机变量相互独立性的概念。

【重难点提示】随机变量的相互独立性。对分布函数的理解及用该函数求概率问题，联合分 布与边缘分布之间的关系。

第一节　多维随机变量及其分布

一、多维离散型随机变量的表达方式

二、多维连续型随机变量的表达方式

第二节　边缘分布

一、联合分布与边缘分布的概念及其关系

二、离散型随机变量边缘分布列

三、连续型随机变量的边缘概率密度函数

第三节　条件分布

一、二维离散型随机变量的两种条件分布的计算方法

二、二维连续型随机变量的两种条件分布概率密度函数的计算

第四节　多维随机变量的特征数

一、了解特征数条件

二、利用随机变量的特征数解决实际问题

第五节　条件分布与条件期望

一、条件分布的定义；

二、条件分布的定义

三、条件期望；

第四章 大数定律与中心极限定理

【知识点提示】了解切比雪夫不等式，切比雪夫大数定律与贝努里大数定律，辛钦大数定律中心极限定理（独立同分布的中心极限定理、李雅普洛夫、棣莫佛 - 拉普拉斯中心极限定理）。

基本知识点：

1、贝努里大数定律的内容与含义。

2、中心极限定理的内容与含义。

【重难点提示】对贝努里大数定律，中心极限定理的理解

第一节　特征函数

一、特征函数的定义；

二、性质

第二节　大数定律

一、切比雪夫不等式及其意义

二、切比雪夫大数定律与贝努里大数定律的内容与含义。

第三节 随机变量序列的两种收敛性

一、收敛行描述；

二、性质

第四节 中心极限定理

一、独立同分布的中心极限定理的内容与含义

二、李雅普洛夫的内容与含义

三、棣莫佛 - 拉普拉斯中心极限定理的内容与含义。

2024 年北京大学 432 统计学考研核心题库

《概率论与数理统计教程》考研核心题库之解答题精编

1. 据以往资料表明，某一 3 口之家，患某种传染病有以下规律：P{孩子得病}=0.6，P{母亲得病|孩子得病}=0.5；P{父亲得病|母亲及孩子得病}=0.4。求母亲及孩子得病但父亲未得病的概率。

【答案】设“母亲得病”为事件 A，“孩子得病”为事件 B，“父亲得病”为事件 C；

$$P(AB)=P(A|B)\cdot P(B)=0.5\times 0.6=0.3;$$

$$P(C|AB)=\frac{P(ABC)}{P(AB)}\Rightarrow P(ABC)=P(C|AB)\cdot P(AB)=0.12$$

$$P(AB\overline{C})=P(AB)-P(ABC)=0.18。$$

2. 尽管在几何教科书中已经讲过用圆规和直尺三等分一个任意角是不可能的，但每年总有一些“发明者”撰写关于用圆规和直尺将角三等分的文章，设某地区每年撰写此类文章的篇数 X 服从参数为 6 的泊松分布，求明年没有此类文章的概率。

【答案】明年没有此类文章的概率

$$p\{X=0\}=6^0\frac{e^{-6}}{0!}=0。0025$$

3. 设二维随机变量 (X,Y) 的密度函数：$f(x,y)=\begin{cases}A, & 0<x<2,|y|<x\\ 0, & 其他\end{cases}$

(1)求常数 A 的值；(2)求边缘概率密度 $f_X(x),f_Y(y)$；

(3) X 和 Y 是否独立?

【答案】(1)由 $\int_{-\infty}^{+\infty}f(x,y)dy=1$，得 $A=1/4$

(2) $f_X(x)=\int_{-\infty}^{+\infty}f(x,y)dy=\begin{cases}\int_{-x}^{x}1/4dy, & 0\le x\le 2\\ 0, & 其他\end{cases}=\begin{cases}x/2, & 0\le x\le 2\\ 0, & 其他\end{cases}$

$$f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)dx=\begin{cases}\int_{-y}^{2}1/4dx, & -2\le y<0\\ \int_{y}^{2}1/4dx, & 0\le y<2\\ 0, & 其他\end{cases}=\begin{cases}(2+y)/4, & -2\le y<0\\ (2-y)/4, & 0\le y<2\\ 0, & 其他\end{cases}$$

(3) $f_X(x)f_Y(y)\ne f(x,y)$，不独立

4. 设 A、B、C 为三事件，且 $P(A)=P(B)=P(C)=1/4$，$P(AB)=P(BC)=0$，$P(AC)=1/8$，求 A、B、C 至少有一个发生的概率。

【答案】至少有一个发生的概率为 $P(A\cup B\cup C)$，因 $P(AB)=P(BC)=0$，所以 $AB=BC=\Phi$，$(A\cup C)B=\Phi$，则

$P(A\cup B\cup C)=P(A\cup C\cup B)=P(A\cup C)+P(B)=P(A)+P(C)-P(AC)$

$+P(B)=\frac{1}{4}+\frac{1}{4}-\frac{1}{8}+\frac{1}{4}=\frac{5}{8}$。

5. 设一群人中有 37.5%的人的血型为 A 型，20.9%为 B 型，33.7%为 O 型，7.9%为 AB 型，已知能允许输血的血型配对如下表,现在在一群人中任选一人为输血者,再任选一人为需要输血者,问输血能成功的概率是多少?

输血者 \ 受血者	A 型	B 型	AB 型	O 型
A 型	√	×	√	√
B 型	×	√	√	√
AB 型	√	√	√	√
O 型	×	×	×	√

其中: √表示允许输血；×表示不允许输血。

【答案】设 A_1,A_2,A_3,A_4 分别表示 A,B,AB,O 型输血；B_1,B_2,B_3,B_4 分别表示 A,B,AB,O 型受血，则

$P(\text{输血成功})=P[(A_1\cup A_2\cup A_3\cup A_4)\cap(B_1\cup B_2\cup B_3\cup B_4)]$

$=P(A_1)\cdot P(B_1)+P(A_1)P(B_3)+P(A_2)P(B_2)+P(A_2)P(B_3)+P(A_3)P(B_1)$

$+P(A_3)P(B_2)+P(A_3)P(B_3)+P(A_4)$=61.98%(利用 A_i,B_j 的独立性, $A_iB_j=\Phi$)

6. 将两信息分别编码为 A 和 B 传递出去，接收站收到时，信息 A 被误收作 B 的概率为 0.02，而 B 被误收作 A 的概率为 0.01.信息 A 与信息 B 传送的频繁程度为 2∶1.若接收站收到的信息是 A，问原发信息是 A 的概率是多少?

【答案】令 $H=\{\text{原发信息是 A}\}$，$C=\{\text{收到的信息是 A}\}$，则

$$P(H|C)=\frac{P(H)P(C|H)}{P(H)P(C|H)+P(\bar{H})P(C|\bar{H})}=\frac{\frac{2}{3}\times 0.98}{\frac{2}{3}\times 0.98+\frac{1}{3}\times 0.01}=\frac{196}{197}=0.995$$

7. 求标准正态分布上 a 分为点，(1)a=0.01,求 z_a;(2)a=0.003,求 $z_a, z_{\frac{a}{2}}$

【答案】(1)由 a 分为点的定义，则应查 1-a=0.99 得 $\Phi(2.33)=0.99, z_{0.01}$=2.33

(2) $\Phi(2.75)=0.997,\Phi(2.96)=0.9985,z_{0.003}=2.75,z_{0.0015}$=2.96

8. **(1)由统计物理学知，分子运动速度的绝对值 X 服从马克思韦尔分布，其概率密度为**

$$f(x)=\begin{cases}Ax^2e^{-\frac{x^2}{b}}, x>0\\0\cdots\cdots\cdots, x\le 0\end{cases}$$

其中$b=\frac{m}{2kT}$,k为常数，T 为绝对温度，m 是分子的质量,试确定常数 A。

(2)研究了英格兰在 1875 年——1951 年期间，在矿山发生导致 10 人或 10 人以上死亡的事故的频繁程度，得到相继两次事故之间的时间 T(以日计算)服从指数分布，其概率密度为

$$f(x)=\begin{cases}\frac{1}{241}e^{-\frac{x}{241}}, x>0\\0\cdots\cdots\cdots, x\le 0\end{cases}$$

求分布函数 F(x)，并求概率 p{50<T<100}。

【答案】(1) $\int_0^{\infty}e^{-x^2}dx=\frac{1}{2}\int_{-\infty}^{\infty}e^{-x^2}dx=\frac{\sqrt{\pi}}{2}$

由$\int_{-\infty}^{\infty}f(x)dx=\int_0^{\infty}Ax^2e^{-\frac{x^2}{b}}dx=1$和上式易知

$$A=\frac{4}{b\sqrt{b\pi}}$$

(2)利用积分可求得$F(x)=\begin{cases}0\cdots\cdots\cdots, x\le 0\\1-e^{-\frac{x}{241}}, x>0\end{cases}$

p{50<T<100}=F(100)-F(50)

$=e^{-\frac{50}{241}}-e^{-\frac{100}{241}}$

9. **设A,B是两事件，且$P(A)=0.6,P(B)=0.7$.问：**

(1)在什么条件下$P(AB)$取到最大值，最大值是多少？

(2)在什么条件下$P(AB)$取到最小值，最小值是多少？

【答案】由于$P(AB)=P(A)+P(B)-P(A\cup B)$，所以

(1)当$P(A\cup B)=0.7$时，$P(AB)$取最大值 0.6;

(2)当$P(A\cup B)=1$时，$P(AB)$取最小值 0.3.

10. **设一批木材的小头直径$X\sim N(\mu,\sigma^2)$,(单位 cm),今抽出 12 根测得$\overline{x}=11.2$,$s^2=1.44$.**

(1)若已知$\sigma=1.2$,求μ的置信水平为95%的区间估计.

(2)若σ未知,$\mu\ge 12$为合格,问该批木材是否合格?取显著性水平$\alpha=0.05$,

即假设检验: $H_0:\mu\ge 12,\ H_1:\mu<12.$

【答案】(1) $\sigma=1.2$, $n=12,1-\alpha=95\%,\dfrac{\alpha}{2}=0.025,u_{1-\alpha/2}=u_{0.975}=1.96$

μ 的置信区间为 $(\overline{X}\pm\dfrac{\sigma}{\sqrt{n}}u_{1-\alpha/2})=(11.2\pm\dfrac{1.2}{\sqrt{12}}\cdot 1.96)=(10.5210,11.8790)$

(2)假设 $H_0:\mu\le 12$,， $H_1:\mu>12.$

σ^2 未知，检验统计量为 $t=\dfrac{\overline{X}-\mu_0}{S/\sqrt{n}}$， $\mu_0=12,n=12$

拒绝域为 $t\ge t_{1-\alpha}(n-1)=t_{0.95}(n-1)=t_{0.95}(11)=1.7959$

又 $\bar{x}=11.2,s=1.2$, $t=\dfrac{11.2-12}{1.2/\sqrt{12}}=-2.3094<t_{0.95}(11)=1.7959$,

t 没落在拒绝域内，接受 H_0 ,即认为该批木材不合格.

11．设 X 和 Y 是两个相互独立的随机变量， $X\sim U[0,0.2]$，Y 的概率密度函数为

$$f_Y(y)=\begin{cases}5e^{-5y}, & y>0,\\ 0, & y\le 0.\end{cases}$$

（1）求 X 和 Y 的联合概率密度 $f(x,y)$；（2）求概率 $P(Y\le X)$.

【答案】（1）由题意知， X 的概率密度函数为

$$f_X(x)=\begin{cases}5, & 0<x<0.2;\\ 0, & \text{其它}.\end{cases}$$

因为 X 和 Y 相互独立，故 X 和 Y 的联合概率密度

$$f(x,y)=f_X(x)f_Y(y)=\begin{cases}25e^{-5y}, & 0<x<0.2,\ y>0;\\ 0, & \text{其它}.\end{cases}$$

$$P(Y\le X)=\iint\limits_{y\le x}f(x,y)dxdy=\int_0^{0.2}dx\int_0^x 25e^{-5y}dy=5\int_0^{0.2}(1-e^{-5x})dx=e^{-1}.$$

12．由某机器生产的螺栓的长度(cm)服从参数为 μ =10.05, $\sigma=0.06$ 的正态分布，规定长度在范围 10.05~10.12 内为合格品。求一螺栓为不合格品的概率。

【答案】设螺栓的长度为 X，一螺栓为不合格品的概率

1-p{10.05-0.12<X<10.05+0.12}=1- $\Phi\left\{\dfrac{0.12}{0.06}\right\}+\Phi\left(\dfrac{-0.12}{0.06}\right)$

=0.0456

2024年北京大学432统计学考研题库[仿真+强化+冲刺]

北京大学432统计学考研仿真五套模拟题

2024年概率论与数理统计教程五套仿真模拟题及详细答案解析（一）

一、计算题

1. **将指数回归模型**

$$Y=\beta_1 X^{\beta_2} e^{\varepsilon}$$

转化为一个线性回归模型(即对数-对数模型)，分析Y与X之间的弹性系数的特点.

【答案】将指数回归模型 $Y=\beta_1 X^{\beta_2} e^{\varepsilon}$ 两边取对数，得到

$$\ln Y=\ln \beta_1+\beta_2 \ln X+\varepsilon$$

即作对数变换 $Z=\ln Y, w=\ln x$，$a=\ln \beta_1$，得线性回归模型

$$Z=a+\beta_2 w+\varepsilon$$

对指数回归模型 $Y=\beta_1 X^{\beta_2} e^{\varepsilon}$，计算Y与X之间的弹性系数

$$E=\frac{X}{Y}\cdot\frac{dY}{dX}=\frac{X}{\beta_1 X^{\beta_2} e^{\varepsilon}}\times\beta_1\beta_2 X^{\beta_2-1} e^{\varepsilon}=\beta_2$$

故弹性系数为 β_2，也就是对数变换后得到的线性回归模型中 w 前的系数.

2. **下面给出了用于计算器的四种类型的电路的响应时间(单位：ms):**

电路类型	响应时间
Ⅰ	19，22，20，18，15
Ⅱ	24，31，27，28
Ⅲ	17，15，17，23，18
Ⅳ	18，23，19

试判断不同类型的电路响应时间是否有显著差异。(显著性水平 $\alpha=0.05$)

【答案】以 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 分别表示四种不同类型电路响应时间的效应，需要检验假设

$$H_0: \alpha_1=\alpha_2=\alpha_3=\alpha_4=0$$

由题设，$r=4$，$n_1=5$，$n_2=4$，$n_3=5$，$n_4=3$，$n=17$，容易计算

$$T_1=94,\ T_2=110,\ T_3=90,\ T_4=60,\ T=354,$$

$$Q=\sum_{i=1}^{r}\sum_{j=1}^{n_i} X_{ij}-\frac{T^2}{n}=342.47,$$

$$Q_A=\sum_{i=1}^{r}\frac{T_i^2}{n}-\frac{T^2}{n}=240.67,$$

$$Q_e=Q-Q_A=101.8$$

得方差分析表如下

方差来源	平方和	自由度	均方	F值
因素	240.67	3	80.22	10.24
误差	101.8	13	7.83	
总和	342.47	16		

由于 $F=10.24>F_{0.05}(3,13)=3.41$，故拒绝 H_0，即认为不同类型的电路响应时间有显著差异。

3. 在总体 N(12，4)中随机抽一容量为 5 的样本X_1,X_2,X_3,X_4,X_5.(1)求样本均值与总体平均值之差的绝对值大于 1 的概率.(2)求概率$P\{\max\{X_1,X_2,X_3,X_4,X_5\}>15\}$.(3)求概率$P\{\min\{X_1,X_2,X_3,X_4,X_5\}<10\}$.

【答案】(1)$\because \overline{X}\sim N\left(12,\frac{4}{5}\right)$，则$\frac{\overline{X}-12}{2/\sqrt{5}}\sim N(0,1)$

$$\therefore P\{|\overline{X}-12|>1\}=P\left\{\left|\frac{\overline{X}-12}{\sqrt{4/5}}\right|>\frac{\sqrt{5}}{2}\right\}$$

$$=2-2\Phi\left(\frac{\sqrt{5}}{2}\right)=2\times[1-\Phi(1.12)]$$

$$=2\times[1-0.8686]=0.2628$$

(2) $P\{\max\{X_1,X_2,X_3,X_4,X_5\}>15\}$

$=1-P\{\max\{X_1,X_2,X_3,X_4,X_5\}\leqslant 15\}$

$=1-P\{X_1\leqslant 15,X_2\leqslant 15,X_3\leqslant 15,X_4\leqslant 15,X_5\leqslant 15\}$

$=1-\prod_{i=1}^{5}P\{X_i\leqslant 15\}=1-\prod_{i=1}^{5}P\left\{\frac{X_i-12}{2}\leqslant\frac{15-12}{2}\right\}$

$=1-[\Phi(1.5)]^5=1-(0.9332)^5=0.2923$

(3) $P\{\min\{X_1,X_2,X_3,X_4,X_5\}<10\}$

$=1-P\{\min\{X_1,X_2,X_3,X_4,X_5\}\geqslant 10\}$

$=1-P\{X_1\geqslant 10,X_2\geqslant 10,X_3\geqslant 10,X_4\geqslant 10,X_5\geqslant 10\}$

$=1-\prod_{i=1}^{5}P\{X_i\geqslant 10\}=1-\prod_{i=1}^{5}P\left\{\frac{X_i-12}{2}\geqslant\frac{10-12}{2}\right\}$

$=1-[1-\Phi(-1)]^5=1-[\Phi(1)]^5=1-(0.8413)^5$

$=0.5785$

4. 假设随机变量X_1,X_2,X_3,X_4相互独立且同分布，

$$P\{X_i=0\}=0.6,\ P\{X_i=1\}=0.4\quad(i=1,2,3,4)$$

求行列式$X=\begin{vmatrix}X_1 & X_2\\ X_3 & X_4\end{vmatrix}$的概率分布.

【答案】由题设，

$$X=\begin{vmatrix}X_1 & X_2\\ X_3 & X_4\end{vmatrix}=X_1X_4-X_2X_3$$

设$Y_1=X_1X_4,Y_2=X_2X_3$.由于已知X_1,X_2,X_3,X_4相互独立且同分布，则Y_1和Y_2也独立同分布，且有

$$P\{Y_1=0\}=P\{Y_2=0\}$$
$$=P\{X_1=0,X_4=0\}+P\{X_1=0,X_4=1\}+P\{X_1=1,X_4=0\}$$
$$=0.36+0.24\times 2=0.84,$$

从而

$$P\{Y_1=1\}=P\{Y_2=1\}=1-P\{Y_1=0\}=1-0.84=0.16,$$

则

$$P\{X=-1\}=P\{Y_1-Y_2=-1\}=P\{Y_1=0,Y_2=1\}$$
$$=0.84\times 0.16=0.1344,$$
$$P\{X=1\}=P\{Y_1-Y_2=1\}=0.84\times 0.16=0.1344,$$

因此

$$P\{X=0\}=1-0.1344-0.1344=0.7312.$$

综合得，X 的概率分布为$X\sim\begin{pmatrix}-1 & 1 & 0\\ 0.1344 & 0.1344 & 0.7312\end{pmatrix}$

5. **设系统由100个相互独立的部件组成，运行期间每个部件损坏的概率为0.1，利用切比雪夫不等式估计部件是完好的个数在85个和95个之间的概率.**

【答案】设X表示部件是完好的个数，则X～B(100,0.9),得

$$E(X)=np=100\times0.9=90, D(X)=np(1-p)=100\times0.9\times0.1=9$$

由切比雪夫不等式得$P\{85<X<95\}=P\{|X-90|<5\}\geqslant1-\frac{9}{25}=\frac{16}{25}=0.64$.

则部件是完好的个数在85个和95个之间的概率大于等于0.64.

6. **设总体X服从$N(\mu,\sigma^2)$，从该总体中抽取简单随机样本$X_1,X_2,\cdots,X_{2n}(n\geqslant2)$，其样本均值为$\bar{X}=\frac{1}{2n}\sum_{i=1}^{2n}X_i$，求统计量$Y=\sum_{i=1}^{n}(X_i+X_{n+i}-2\bar{X})^2$的数学期望E(Y).**

【答案】令$Y_i=X_i+X_{n+i}$，则$Y_i\sim N(2\mu,2\sigma^2)$，其样本均值为

$$\frac{1}{n}\sum_{i=1}^{n}(X_i+X_{n+i})+\frac{1}{n}\sum_{i=1}^{2n}X_i=2\bar{X}$$

样本方差为$\frac{1}{n-1}\sum_{i=1}^{n}(Y_i-2\bar{X})^2=\frac{1}{n-1}Y$

$$\because E\left(\frac{1}{n-1}Y\right)=2\sigma^2$$

$$\therefore E(Y)=2(n-1)\sigma^2$$

7. **设随机变量(X,Y)的概率密度为：**

$$f(x,y)=\begin{cases}k(6-x-y), & 0<x<2,2<y<4,\\ 0, & \text{其他}.\end{cases}$$

(1)确定常数k.

(2)求$P\{X<1,Y<3\}$.

(3)求$P\{X<1.5\}$.

(4)求$P\{X+Y\leqslant4\}$.

【答案】(1)

$$\because\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x\mathrm{d}y=\int_0^2\mathrm{d}x\int_2^4k(6-x-y)\mathrm{d}y$$

$$=k\int_0^2(6-2x)\mathrm{d}x=8k=1$$

$$\therefore k=\frac{1}{8}$$

(2)

$$P\{X<1,Y<3\}=\int_0^1\mathrm{d}x\int_2^3\frac{1}{8}(6-x-y)\mathrm{d}y$$

$$=\frac{1}{8}\int_0^1\left[(6-x)-\frac{5}{2}\right]\mathrm{d}x$$

$$=\frac{1}{8}\times\left(\frac{7}{2}-\frac{1}{2}\right)=\frac{3}{8}$$

(3)

$$P\{X<1.5\}=\int_{-\infty}^{1.5}\left[\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}y\right]\mathrm{d}x$$

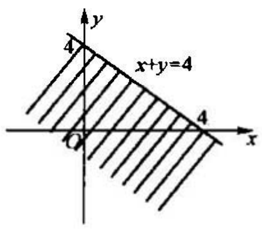
$$=\int_0^{1.5}\left[\int_2^4\frac{1}{8}(6-x-y)\mathrm{d}y\right]\mathrm{d}x$$

$$=\frac{1}{8}\int_0^{1.5}[2(6-x)-6]\mathrm{d}x$$

$$=\frac{1}{8}\int_0^{1.5}(6-2x)\mathrm{d}x$$

$$=\frac{1}{8}[6\times 1.5-(1.5)^2]=\frac{27}{32}$$

(4)将(X，Y)看做是平面上随机点的坐标，即有$\{X+Y\leqslant 4\}=\{(X,Y)\in D\}$，其中 D 为 XOY 平面上直线 x+y=4 下方的部分.



图

$$P\{X+Y\leqslant 4\}=P\{(X,Y)\in D\}$$

$$=\iint_D f(x,y)\mathrm{d}x\mathrm{d}y$$

$$=\int_0^2\mathrm{d}x\int_2^{4-x}\frac{1}{8}(6-x-y)\mathrm{d}y$$

$$=\frac{1}{8}\int_0^2\left[(6-x)(2-x)-\frac{(4-x)^2-4}{2}\right]\mathrm{d}x$$

$$=\frac{1}{16}\int_0^2(12-8x+x^2)\mathrm{d}x=\frac{2}{3}$$

8. 槲寄生是一种寄生在大树上部树枝上的寄生植物.它喜欢寄生在年轻的大树上.下面给出在一定条件下完成的试验中采集的数据.

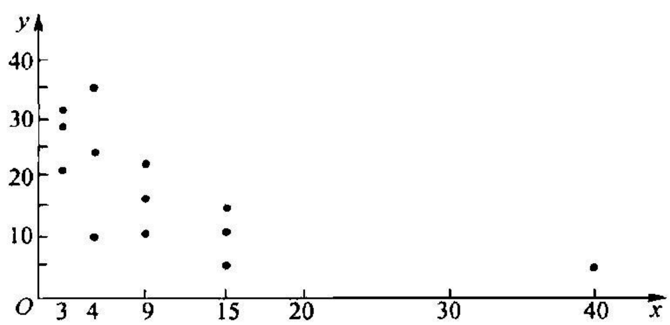
x:大树的年龄(年)	3	4	9	15	40
y:每株大树上槲寄生的株数	28	10	15	6	1
	33	36	22	14	1
	22	24	10	9	

(1)作出(x_i,y_i)的散点图.

(2)令$z_i=\ln y_i$，作出(x_i,z_i)的散点图.

(3)以模型$Y=ae^{bx}\varepsilon,\ln\varepsilon\sim N(0,\sigma^2)$拟合数据，其中$a,b,\sigma^2$与 x 无关，试求曲线回归方程$\hat{y}=\hat{a}\exp(\hat{b}x)$

【答案】(1)作(x_i,y_i)的散点图如下图所示



(2)n=15，令$z=\ln y,\alpha=\ln a$，模型化为 z=a+bx.所需计算列表如下：