厦门大学《概率论与数理统计》课程期末试卷

信息学院信息与通信工程系 19 级计算机类专业

学年学期：2019-2020 学年春季学期

主考教师：王 琳　　试卷类型：B 卷(　)C 卷(√)

一、 选择题（在各小题的四个备选答案中选出一个正确答案，填在题后的括号中，本大题共 5 个小题，每小题 3 分，总计 15 分）

1. 设 A,B,C 为三个事件，用 A,B,C 的运算关系表示“三个事件恰好一个发生”为（　　）。

A. $A\cup B\cup C$　　B. $A\overline{B}\overline{C}+\overline{A}B\overline{C}+\overline{A}\overline{B}C$

C. $\Omega-ABC$　　D. $\overline{A}\overline{B}\overline{C}+A\overline{B}\overline{C}+\overline{A}B\overline{C}+\overline{A}\overline{B}C$

知识点：随机事件的概念
答案：B

2. 设 $X\sim N(1,2^2)$，$X_1,X_2,\cdots,X_n$ 为 X 的样本，则　　（　　）

A. $\dfrac{\bar{X}-1}{2}\sim N(0,1)$　　B. $\dfrac{\bar{X}-1}{4}\sim N(0,1)$

C. $\dfrac{\bar{X}-1}{\sqrt{2}}\sim N(0,1)$　　D. $\dfrac{\bar{X}-1}{2/\sqrt{n}}\sim N(0,1)$

知识点：抽样分布
答案：D

3. 设随机变量 X 服从正态分布 $N(\mu,\sigma^2)$，则随着 σ 的增大，概率 $P\{|X-\mu|<\sigma\}$（ ）。

A. 单调增大　B. 单调减少　C. 保持不变　D. 增减不定
知识点：正态分布。
答案：C

4. 设X~N(μ,σ^2)，其中 μ 已知，σ^2 未知，X_1,X_2,X_3 为样本，则下列选项中不是统计量的是（）。

A. $X_1+X_2+X_3$　　B. $\max\{X_1,X_2,X_3\}$

C. $\sum_{i=1}^{3}\dfrac{X_i^2}{\sigma^2}$　　D. $X_1-\mu$

知识点：统计量的定义

答案：C

5. 在假设检验中，显著性水平表示(　　　)

A. $P\{接受H_0 \mid H_0不真\}=\alpha$　　　　B. $P\{拒绝H_0 \mid H_0真\}=\alpha$

C. $P\{接受H_0 \mid H_0真\}=\alpha$　　　　D. $P\{拒绝H_0 \mid H_0不真\}=\alpha$

知识点：显著性水平的基本定义

答案：B

二、 计算题（本大题共 5 小题，每小题 15 分，共计 75 分）

1.（1）设一批混合麦种中，一、二、三等品分别占 80%、15%、5%，三个等级的发芽率依次为 0.98、0.95、0.8 求这批麦种的发芽率

（2）设有一箱同类产品是由三家工厂生产得，其中 $\frac{1}{2}$ 是第一家工厂生产的，其余两家各生产 $\frac{1}{4}$。又知第一、二、三家工厂生产的产品分别有 2%、4%、5%的次品。现从箱子种任取一件产品，若取到的是次品，它是第一家工厂生产的概率是多少？

答案：（1）设 $B=\{能发芽\}$，$A_i=\{取的是第i等品\}$，$i=1,2,3$ ····················1 分

$P(A_1)=0.8$，$P(A_2)=0.15$，$P(A_3)=0.05$

$P(B|A_1)=0.98$，$P(B|A_2)=0.95$，$P(B|A_3)=0.8$，································2 分

由全概率公式得：

$P(B)=\sum_{i=1}^{3}P(A_i)P(B|A_i)$ ··3 分

$=0.8\times0.98+0.15\times0.95+0.05\times0.8=0.9665$ ·····································1 分

（2）设 $B=\{取到次品\}$，$A_i=\{取到的产品是第i家工厂生产\}$，$i=1,2,3$ ······1 分

$P(A_1)=0.5$，$P(A_2)=0.25$，$P(A_3)=0.25$

$P(B|A_1)=0.02$，$P(B|A_2)=0.04$，$P(B|A_3)=0.05$ ································3 分

由贝叶斯公式得：

$P(A_1|B)=\frac{P(A_1B)}{P(B)}=\frac{P(A_1)P(B|A_1)}{\sum_{i=1}^{3}P(A_i)P(B|A_i)}$ ···3 分

$$=\frac{0.5\times0.02}{0.5\times0.02+0.25\times0.04+0.25\times0.05}=\frac{4}{13}\approx0.3077$$ ……………………………1 分

2.（1）随机变量(X,Y)的概率密度为

$$f(x,y)=\begin{cases}\frac{e^{-(x+y)}}{1-e^{-1}}, & 0<x<1,0<y<\infty\\ 0, & \text{其他}\end{cases}$$

证明随机变量X和Y的独立性并求函数$U=\max(X,Y)$的分布函数。

（2）某种计算器在进行加法时，将每个加数的小数部分删去，设所有这部分操作导致的误差相互独立且在（0，1）上服从均匀分布。若将 4800 个数相加，运用中心极限定理求：误差总和的超过 2414 的概率是多少？

（$\Phi(0.7)=0.7580,\Phi(0.8)=0.7881,\Phi(0.9)=0.8159$）

答案：

（1）解：

$$f_X(x)=\int_{-\infty}^{\infty}f(x,y)dy=\begin{cases}\frac{1}{1-e^{-1}}\int_0^{\infty}e^{-(x+y)}dy=\frac{e^{-x}}{1-e^{-1}}, & 0<x<1\\ 0, & \text{其他}\end{cases}$$ ………………… **1.5 分**

$$f_Y(y)=\int_{-\infty}^{\infty}f(x,y)dx=\begin{cases}\frac{1}{1-e^{-1}}\int_0^{1}e^{-(x+y)}dy=e^{-y}, & y>0\\ 0, & \text{其他}\end{cases}$$ ………………………… **1.5 分**

由上可知$f(x,y)=f_X(x)f_Y(y)$，故X,Y相互独立，分别计$U=\max(X,Y)$，X和Y的分布函数为$F_U(u),F_X(x)$和$F_Y(y)$，则有

$$F_U(u)=F_X(u)F_Y(u)$$ …………………………………………………………… **1 分**

又

$$F_X(u)=\int_{-\infty}^{u}f_X(x)dx=\begin{cases}0, & u<0\\ \int_0^u\frac{e^{-x}}{1-e^{-1}}dx, & 0\le u<1\\ 1, & u\ge1\end{cases}=\begin{cases}0, & u<0\\ \frac{1-e^{-u}}{1-e^{-1}}, & 0\le u<1\\ 1, & u\ge1\end{cases}$$ …………… **1.5 分**

$$F_Y(u)=\int_{-\infty}^{u}f_Y(y)dy=\begin{cases}0, & u<0\\ \int_0^u e^{-y}dy, & u\ge0\end{cases}=\begin{cases}0, & u<0\\ 1-e^{-u}, & u\ge0\end{cases}$$ ………………………… **1.5 分**

将$F_X(u)$和$F_Y(u)$的表达式代入$F_U(u)=F_X(u)F_Y(u)$，得到$U=\max(X,Y)$的分布函

数为

$$F_U(u)=\begin{cases}0, & u<0\\ \dfrac{(1-e^{-u})^2}{1-e^{-1}}, & 0\le u<1\\ 1-e^{-u}, & u\ge 1\end{cases}$$ ………………………………………………………… **1 分**

（2）解：设第 k 个加数的舍入误差为 $X_k(k=1,2,\ldots,4800)$，已知 X_k 在（0，1）上服从均匀分布，故知 $E(X_k)=0.5, D(X_k)=\dfrac{1}{12}$ …………………………………… **2 分**

记 $X=\sum_{k=1}^{4800}X_k$，由中心极限定理，当 n 充分大时有近似公式

$$P\left\{\frac{\sum_{k=1}^{4800}X_k-4800\times 0.5}{\sqrt{4800}\sqrt{1/12}}\le x\right\}\approx\Phi(x)$$ ………………………………………………………… **2 分**

于是

$$P\{X>2414\}=1-P\{X\le 2414\}$$
$$=1-P\left\{\frac{X-2400}{\sqrt{4800}\sqrt{1/12}}\le\frac{2414-2400}{\sqrt{4800}\sqrt{1/12}}\right\}$$
$$\approx 1-\Phi(\frac{14}{20})=1-\Phi(0.7)$$ …………………………………………… **2 分**

$=1-0.7580=0.2420$ ………………………………………………………………… **1 分**

即误差总和的绝对值超过 2414 的概率近似地为 0.2420.

3. 设总体 X 的概率密度为 $f(x;\sigma)=\dfrac{1}{2\sigma}e^{-\frac{|x|}{\sigma}},-\infty<x<+\infty$，其中 $\sigma\in(0,+\infty)$ 为未知参数，$X_1,X_2,\cdots,X_n$ 为来自总体 X 的简单随机样本。记 σ 的最大似然估计量为 $\hat{\sigma}$。

（1）求 $\hat{\sigma}$；（2）证明：$\hat{\sigma}$ 是 σ 的无偏估计；（3）求 $D\left(\hat{\sigma}\right)$.

答案：

（1） 设 $x_1,x_2,\cdots,x_n$ 为 $X_1,X_2,\cdots,X_n$ 的观测值，

则似然函数为：$L(\sigma)=\prod_{i=1}^{n}\frac{1}{2\sigma}e^{-\frac{|x_i|}{\sigma}}=\frac{1}{(2\sigma)^n}e^{-\frac{\sum_{i=1}^{n}|x_i|}{\sigma}}$ ，…………………………… 3 分

取对数得到：$\ln L(\sigma)=-n\ln 2-n\ln\sigma-\frac{1}{\sigma}\sum_{i=1}^{n}|x_i|$，

则 $\frac{d\ln L(\sigma)}{d\sigma}=-\frac{n}{\sigma}+\frac{1}{\sigma^2}\sum_{i=1}^{n}|x_i|$。

令 $\frac{d\ln L(\sigma)}{d\sigma}=0$，……………………………………………………………… 2 分

解得 σ 的最大似然估计值为：$\hat{\sigma}=\frac{1}{n}\sum_{i=1}^{n}|x_i|$，则 σ 的最大似然估计量为：

$\hat{\sigma}=\frac{1}{n}\sum_{i=1}^{n}|X_i|$。……………………………………………………………… 1 分

（2） 由期望计算公式可得：

$$E\left(\hat{\sigma}\right)=E\left(\frac{1}{n}\sum_{i=1}^{n}|X_i|\right)=\frac{1}{n}E\left(\sum_{i=1}^{n}|X_i|\right)=E\left(|X|\right)$$

$$=\int_{-\infty}^{+\infty}|x|\frac{1}{2\sigma}e^{-\frac{|x|}{\sigma}}dx=\int_0^{+\infty}\frac{x}{\sigma}e^{-\frac{x}{\sigma}}dx=\sigma$$

因此，$\hat{\sigma}$ 是 σ 的无偏估计。……………………………………………………… 3 分

（3） $D\left(\hat{\sigma}\right)=D\left(\frac{1}{n}\sum_{i=1}^{n}|X_i|\right)=\frac{1}{n^2}D\left(\sum_{i=1}^{n}|X_i|\right)=\frac{1}{n}D\left(|X|\right)$，………………… 2 分

而 $E\left(|X|^2\right)=\int_{-\infty}^{+\infty}|x|^2\frac{1}{2\sigma}e^{-\frac{|x|}{\sigma}}dx=\int_0^{+\infty}\frac{x^2}{\sigma}e^{-\frac{x}{\sigma}}dx=\sigma^2\int_0^{+\infty}\left(\frac{x}{\sigma}\right)^2e^{-\frac{x}{\sigma}}d\left(\frac{x}{\sigma}\right)=2\sigma^2$，… 2 分

由（2）可知：$E\left(|X|\right)=\sigma$，

$D\left(|X|\right)=E\left(|X|^2\right)-\left[E\left(|X|\right)\right]^2=2\sigma^2-\sigma^2=\sigma^2$，………………………………… 1 分

因此，$D\left(\hat{\sigma}\right)=\frac{\sigma^2}{n}$.……………………………………………………………… 1 分

4. 设某种清漆的 9 个样品，其干燥时间（以 h 计）分别为

6.0　5.7　5.8　6.5　7.0　6.3　5.6　6.1　5.0

设干燥时间总体服从正态分布 $N\left(\mu,\sigma^2\right)$. 求 μ 的置信水平为 0.95 的置信区间.

（1） 若由以往经验知 $\sigma=0.6(\mathrm{h})$.

（2） 若 σ 为未知.

（已知 $z_{0.025}=1.96, z_{0.05}=1.645, t_{0.025}(8)=2.306, t_{0.05}(8)=1.860$）

解：算得 $\bar{x}=6$， $s^2=0.33$，…………………………………………………………1 分

总体 $X \sim N\left(\mu,\sigma^2\right)$， μ **未知，** σ^2 **已知，** $X_1,X_2,\cdots,X_n$ **是来自 X 的样本, 需要求** μ **的置信水平为** $1-\alpha$ **的置信区间，既需确定随机区间** $(\underline{\theta},\bar{\theta})$, **使得**

$$P\{\underline{\theta}<\mu<\bar{\theta}\}=1-\alpha$$

因为 $\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$，**有**

$$P\left\{-z_{\alpha/2}<\frac{\bar{X}-\sigma}{\sigma/\sqrt{n}}<z_{\alpha/2}\right\}=1-\alpha$$ ……………………………………………………3 分

在上述式子中解出 μ**，得**

$$P\left\{\bar{X}-\frac{\sigma}{\sqrt{n}}z_{\alpha/2}<\mu<\bar{X}+\frac{\sigma}{\sqrt{n}}z_{\alpha/2}\right\}=1-\alpha$$

即得到的一个置信水平为 $1-\alpha$ **的置信区间为**

$$\left(\bar{X}-\frac{\sigma}{\sqrt{n}}z_{\alpha/2},\ \bar{X}+\frac{\sigma}{\sqrt{n}}z_{\alpha/2}\right)$$ ……………………………………………………………3 分

令 $n=9,\sigma=0.6,1-\alpha=0.95,\alpha/2=0.025,z_{0.025}=1.96,$ **并算得** $\bar{x}=6$**，得到** μ **的一个置信水平为 0.95 的置信区间为** $\left(6-\dfrac{0.6}{3}z_{0.025},6-\dfrac{0.6}{3}z_{0.025}\right)=\left(5.608,6.392\right)$. …1 分

（2） σ **未知，由于** $\dfrac{\bar{X}-\mu}{S/\sqrt{n}} \sim t(\mathrm{n}-1)$

有 $P\left\{-t_{\alpha/2}(\mathrm{n}-1)<\dfrac{\bar{X}-u}{S/\sqrt{n}}<t_{\alpha/2}(\mathrm{n}-1)\right\}=1-\alpha$ ………………………………………3 分

在上述式子中解出 μ**，得**

$$P\left\{\bar{X}-\frac{S}{\sqrt{n}}t_{\alpha/2}(\mathrm{n}-1)<\mu<\bar{X}+\frac{S}{\sqrt{n}}t_{\alpha/2}(\mathrm{n}-1)\right\}=1-\alpha$$

即得到的一个置信水平为 $1-\alpha$ 的置信区间为

$$\left(\bar{X}-\frac{S}{\sqrt{n}}t_{\alpha/2}(\mathrm{n}-1),\ \bar{X}+\frac{S}{\sqrt{n}}t_{\alpha/2}(\mathrm{n}-1)\right)$$ ……………………………………………3 分

令 $n=9, 1-\alpha=0.95, \alpha/2=0.025, t_{0.025}(8)=2.306$**，并算得** $\bar{x}=6$**，** $s^2=0.33$**，得到** μ **的一个置信水平为 0.95 的置信区间为**

$$\left(6-\frac{\sqrt{0.33}}{3}\times 2.306, 6-\frac{\sqrt{0.33}}{3}\times 2.306\right)=(5.558, 6.442).$$ …………………………1 分

5.（1）某机器生产的零件长度（单位：cm）$X\sim N(\mu,\sigma^2)$，今抽取容量为 16 的样本，测得样本均值 $\bar{X}=10$，样本方差 $S^2=0.16$。规定该零件平均长度不能超过 9.8cm，问在显著性水平 0.05 下，这批零件是否合格？

（2）某种导线，要求其电阻的标准差不得超过 0.005Ω，今在生产的一批导线中取样品 10 根，测得 $s=0.006\Omega$， 设总体为正态分布，参数均未知，问在显著性水平 $\alpha=0.05$ 下能否认为这批导线的标准差显著地偏大？

（$t_{0.05}(16)=1.746$，$t_{0.05}(15)=1.753$，$t_{0.025}(15)=2.132$；

$\chi^2_{0.05}(8)=15.507$，$\chi^2_{0.05}(9)=16.919$，$\chi^2_{0.05}(10)=18.307$，$\chi^2_{0.05}(11)=19.675$）

答案：

（1）设 $H_0:\mu\le 9.8$，$H_1:\mu>9.8$ …………………………………………………2 分

采用 t 检验法，取检验统计量 $\mathrm{t}=\dfrac{\bar{X}-\mu_0}{S/\sqrt{n}}$，

令 $n=16$，$S^2=0.16$，$\bar{X}=10$，$\alpha=0.05$**，则** $t_\alpha(n-1)=t_{0.05}(15)=1.753$**. 拒绝域为**

$t\ge t_\alpha(n-1)=1.753$ ……………………………………………………………3 分

因 t **的观察值** $t=\dfrac{(10-9.8)}{0.4}\sqrt{16}=2>1.753$**，落在拒绝域内**……………………2 分

故在水平 $\alpha=0.05$ **下拒绝** H_0**，因此可以认为这批零件不合格。**……………1 分

（2）设 $H_0:\sigma\le 0.005, H_1:\sigma>0.005$ ……………………………………………2 分

采用 χ^2 检验，取检验统计量为 $\chi^2=\dfrac{(n-1)S^2}{\sigma_0^2}$，

今 $n=10, s^2=0.006^2, \alpha=0.05$， $\chi_\alpha^2(n-1)=\chi_{0.05}^2(9)=16.919$，拒绝域为

$\chi^2=\dfrac{(n-1)s^2}{\sigma_0^2}\ge\chi_\alpha^2(n-1)=16.919$ ……………………………………………2 分

因 χ^2 的观察值 $\chi^2=\dfrac{9\times(0.006)^2}{0.005^2}=12.96<16.919$， 没有落在拒绝域内………2 分

故在显著性水平 $\alpha=0.05$ 下接受 H_0，

即不能认为这批导线的标准差显著地偏大。 ………………………………………1 分

三、 证明题（本大题共 1 小题，每小题 10 分，共计 10 分）

设 $X_1,X_2,\cdots X_n$ 是独立同分布的随机变量，其数学期望 $E(X_k)=\mu$ 和方差 $D(X_k)=\sigma^2$ 存在且有限，令 $Y_n=\dfrac{X_n+X_{n-1}}{2}, n=1,2,\cdots$，证明：$Y_1,Y_2,\cdots Y_n$ 服从大数定律.

答案：$\because X_1, X_2,\cdots,X_n$ 是独立同分布的随机变量

$\therefore Y_1, Y_2,\cdots,Y_n$ 是独立同分布的随机变量………………………………………1 分

$\therefore E(Y_n)=E(\dfrac{X_n+X_{n-1}}{2})=\dfrac{E(X_n)+E(X_{n-1})}{2}=\mu$

$D(Y_n)=D(\dfrac{X_n+X_{n-1}}{2})=\dfrac{D(X_n)+D(X_{n-1})}{4}=\dfrac{1}{4}\sigma^2$ ……………………………………3 分

$\therefore E(\dfrac{1}{n}\sum_{k=1}^{n}Y_k)=\dfrac{1}{n}\cdot n\cdot\mu=\mu,$

$D(\dfrac{1}{n}\sum_{k=1}^{n}Y_k)=\dfrac{1}{n^2}\cdot\dfrac{n}{2}\cdot\sigma^2=\dfrac{\sigma^2}{2n}$ ……………………………………………………3 分

由切比雪夫不等式得：

$$1 \geq P\left\{\left|\frac{1}{n}\sum_{k=1}^{n} Y_k - \mu\right| < \varepsilon\right\} \geq 1 - \frac{\sigma^2}{2n \cdot \varepsilon^2}$$

$$\Rightarrow \lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{k=1}^{n} Y_k - \mu\right| < \varepsilon\right\} = 1$$

……………………………………………………3 分

$\therefore Y_1$，$Y_2, \cdots, Y_n$ 服从大数定律