厦门大学《概率统计I》课程期末考试试卷

________学院______系______年级______专业

主考教师：______ 试卷类型：（A 卷）

1.

分数	阅卷人

(12分) 设随机变量$X_1,\cdots,X_n,\cdots,$ 相互独立且同分布于参数为λ 的泊松分布，$X_1\sim\pi(\lambda)$。则 $\frac{1}{n}\sum_{k=1}^n X_k$ 和 $\frac{1}{n}\sum_{k=1}^n X_k^2$ 分别依概率收敛吗？如果依概率收敛，分别收敛于什么？

简略证明提示：$\mathbb{E}(X_1)=\lambda, \mathbb{E}\left(X_1^2\right)=\lambda+\lambda^2$。根据大数定律，有 $\frac{1}{n}\sum_{k=1}^n X_k \xrightarrow{\mathbb{P}} \lambda$ 和$\frac{1}{n}\sum_{k=1}^n X_k^2 \xrightarrow{\mathbb{P}} \lambda+\lambda^2$。

2.

分数	阅卷人

(12分) 某保险公司承接了10000辆汽车保险，每辆车每年付228元保险费。若汽车丢失，则车主得赔偿10000 元。假设汽车的丢失率为0.02。问：

(i) 保险公司亏本的概率有多大？

(ii) 其他条件不变，为使保险公司一年的利润不少于2000000元的概率不小于95%，赔偿金至多可设为多少？

简略证明提示：设X为年丢失车数，Y为每辆车的年保费，保险公司的年收入为$10000Y-10000X$。

$$
\begin{aligned}
\mathbb{P}(10000\times 228-10000X<0) &= \mathbb{P}\left(\frac{X-10000\times 0.02}{\sqrt{10000\times 0.02\times 0.98}} > \frac{228-10000\times 0.02}{\sqrt{10000\times 0.02\times 0.98}}\right)\\
&= \mathbb{P}\left(\frac{X-10000\times 0.02}{\sqrt{10000\times 0.02\times 0.98}} > 2\right)\\
&\approx 1-\Phi(2)\\
&\approx 1-0.977=0.023
\end{aligned}
$$

设最大赔偿额度为a

$$
\begin{aligned}
95\% \le \mathbb{P}(10000\times 228-aX>2000000) &= \mathbb{P}\left(\frac{X-10000\times 0.02}{\sqrt{10000\times 0.02\times 0.98}} < \frac{\frac{10000\times 228-2000000}{a}-10000\times 0.02}{\sqrt{10000\times 0.02\times 0.98}}\right)\\
&= \mathbb{P}\left(\frac{X-10000\times 0.02}{\sqrt{10000\times 0.02\times 0.98}} < \frac{\frac{10000\times 28}{a}-200}{14}\right)\\
&\approx \Phi\left(\frac{\frac{10000\times 28}{a}-200}{14}\right)
\end{aligned}
$$

$\frac{\frac{10000\times 28}{a}-200}{14} \geq 1.65 \Leftrightarrow \frac{10000\times 28}{a} \geq 200+14\times 1.65 \Leftrightarrow a \leq \frac{10000\times 28}{200+14\times 1.65} \approx 1255$。

3.

分数	阅卷人

(12分) 设总体X服从$[a,b]$上的均匀分布，$a,b\in\mathbb{R}$未知。$X_1,\cdots,X_n\,(n\in\mathbb{N}_+)$是来自$X$的样本，试求$a,b$的矩估计和极大似然估计。

简略证明提示： $\mu_1=\mathbb{E}(X)=\frac{a+b}{2}$，$\mu_2=\mathbb{E}(X^2)=\frac{a^2+ab+b^2}{3}$。于是$a=\mu_1-\sqrt{3(\mu_2-\mu_1^2)}$，$b=\mu_1+\sqrt{3(\mu_2-\mu_1^2)}$。从而我们得知$a,b$ 的矩法估计量分别为$\overline{X}-\sqrt{\frac{3}{n}\sum_{i=1}^n(X_i-\overline{X})^2}$和$\overline{X}+\sqrt{\frac{3}{n}\sum_{i=1}^n(X_i-\overline{X})^2}$。

$(X_1,\cdots,X_n)$的联合分布密度函数为 $f(x_1,\cdots,x_n)=\left(\frac{1}{b-a}\right)^n,\quad a\leq x_1<\cdots<x_n\leq b.$ 由极大似然估计的思想，我们知道a,b 的矩法估计量分别为$\min_{1\leq i\leq n}X_i$和$\max_{1\leq i\leq n}X_i$。

4.

分数	阅卷人

(12分) 已知幼儿身高服从正态分布$N(\mu,\sigma^2)$，现从5-6岁的幼儿中随机地抽查了10人，其高度（单位：cm）分别为：$115,120,131,115,109,115,115,105,110,112$。求总体均值$\mu$置信度为95%置信区间及标准差$\sigma$的置信度为95%单侧置信下限。

简略证明提示： $\overline{X}=104.7, S^2=50. \frac{\overline{X}-\mu}{S/\sqrt{10}}\sim t(9)$，于是$\mu$的置信度为95%的置信区间为$\left(\overline{X}-S/\sqrt{10}t_{0.025}(9),\overline{X}+S/\sqrt{10}t_{0.025}(9)\right)=(109.64,119.76)$.

我们有$\frac{9S^2}{\sigma^2}\sim\chi^2(9)$，于是$\sigma$的置信度为95%的单侧置信下界为$\sqrt{\frac{9S^2}{\chi^2_{0.05}(9)}}=5.16$.

5.

分数	阅卷人

(12分) 今有某种产品两批，它们分别是**I**,**II**两个工厂所生产的。随机抽取**I**生产的产品10件，测得样本均值为$\bar{x}=76.23(hour)$, 样本方差为$s_1^2=3.32(hour^2)$；抽取**II**生产的产品10件，测得样本均值为$\bar{y}=79.43(hour)$, 样本方差为$s_2^2=2.23(hour^2)$；设两样本独立，且产品的使用寿命分别服从正态总体$N(\mu_1,\sigma_1^2)$和$N(\mu_2,\sigma_2^2)$，$\mu_1,\mu_2,\sigma_1^2,\sigma_2^2$均未知。试检验，相比工厂**I**，工厂**II**生产的产品寿命是否有显著提高（取显著性水平为0.05）？

简略证明提示：

$$H_{01}:\sigma_1^2-\sigma_2^2=0,\quad H_{11}:\sigma_1^2-\sigma_2^2\neq 0.$$

$$\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}\sim F(10-1,10-1).$$

$$F_{0.975}(9,9)\leq\frac{S_1^2}{S_2^2}=1.49\leq F_{0.025}(9,9).$$

接受原假设，可以认为方差相等。

$$H_{02}:\mu_1-\mu_2=0,\quad H_{12}:\mu_1-\mu_2<0.$$

$$T=\frac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{S_\omega\sqrt{\frac{1}{10}+\frac{1}{10}}}\sim t(10+10-2),\quad S_\omega^2=\frac{9S_1^2+9S_2^2}{10+10-2}.$$

$S_w=1.666, t_0=4.3>t_{0.05}(18)=1.7341$

拒绝原假设，认为产品寿命有显著提高。

6.

分数	阅卷人

(12分) 某银行为优化服务柜台数，研究了自2019年1月1日直至2019年1月16日共计22310分钟内该银行相邻两顾客的到达时间间隔数据，统计如下（X表示相邻两顾客的到达时间间隔（单位：分钟），Y表示出现的频数（单位：百人））：

X	0-4	4-9	9-14	14-19	19-24	24-29	29-34	34-39	≥ 39
Y	500	310	260	170	100	80	60	60	80

试检验相邻两顾客的到达时间间隔X是否服从指数分布（取显著性水平为0.05）。

简略证明提示：

$$H_0: f_X(x)=\frac{1}{\theta}e^{-\frac{x}{\theta}}1_{x>0}.$$

θ的极大似然估计为$\overline{X}=\frac{22310}{1620}=13.77$。

A_i	f_i	$\hat{p}_i$	$n\hat{p}_i$	$\frac{f_i^2}{\hat{p}_i}$
$X\in[0,4.5)$	500	0.2788	451.656	553.519
$X\in[4.5,9.5)$	310	0.2196	355.752	270.132
$X\in[9.5,14.5)$	260	0.1527	247.374	273.270
$X\in[14.5,19.5)$	170	0.1062	172.044	167.980
$X\in[19.5,24.5)$	100	0.0739	119.718	83.530
$X\in[24.5,29.5)$	80	0.0514	83.268	76.860
$X\in[29.5,34.5)$	60	0.0358	57.996	62.073
$X\in[34.5,\infty)$	140	0.0816	132.192	148.269
				$\sum=1635.633$

采用统计量的观测值

$$\chi^2=\sum_{i=1}^{k}\frac{f_i^2}{n\hat{p}_i}-n=1635.633-1620=15.633>\chi^2_{0.05}(8-1-1)=12.592,$$

从而认为X并非服从指数分布。

7.

分数	阅卷人

(12分) 某公司专业生产电脑显示器，通过使用3种不同的芯片**I**、**II**、**III**，得出显示器的某个关键参数如下表，每种芯片取5个观测值。

芯片**I**	芯片**II**	芯片**III**
236	257	258
238	253	264
248	255	259
245	254	267
243	261	262

试推断不同的芯片是否对参数具有显著影响（取显著性水平为0.05）。

简略证明提示：

$$S_T=\sum\sum X_{ij}^2-\frac{(\sum\sum X_{ij})^2}{n}=1245.33,\quad S_A=\sum_{j=1}^{3}\frac{(\sum_{i=1}^{5}X_{ij})^2}{5}-\frac{(\sum\sum X_{ij})^2}{n}=1053.33,$$

$$S_E=S_T-S_A=192,$$

F比为

$$\frac{S_A/(3-1)}{S_E/(15-3)}=32.9166>F_{0.05}(3-1,15-3)=3.89,$$

从而说明不同的芯片对参数具有显著影响。

8.

分数	阅卷人

(16分) 某种动物的体测数据中肺活量Y和体重X有密切关系。某实验室对10只该动物进行体测，体测项目包括体重（单位：千克）和肺活量（单位：升），数据见下：

x	3	4	5	7	9	10	12	15	17	20
y	40.5	39.5	41.0	41.5	43.0	42.0	45.0	47.5	53.0	56.0

（1）利用一元线性回归分析确定回归方程$\mu(x)=a+bx$；（2）求误差方差的估计；（3）检验回归效果是否显著（取显著性水平$\alpha=0.05$)；（4）求回归系数b的95%置信区间。

简略证明提示：

$$\sum_i y_i=449,\quad \sum_i x_i=102,\quad \sum_i x_i^2=1338,\quad \sum_i x_iy_i=4855.5,\quad S_{xx}=297.6,\quad S_{xy}=275.7,$$

$$\hat{b}=\frac{S_{xy}}{S_{xx}}=0.926411,\quad \hat{a}=\bar{y}-\bar{x}\hat{b}=44.9-10.2\times0.926411=35.4506,\quad \mu(x)=35.4506+0.926411x,$$

$S_{yy}=282.9$，$Q_e=S_{yy}-\hat{b}S_{xy}=282.9-0.926411\times275.7=27.4885$。$\sigma^2$的无偏估计为

$$\hat{\sigma}^2=\frac{Q_e}{10-2}=3.4361.$$

检验：

$$H_0:b=0,\quad H_1:b\neq0.$$

$$|t|=\frac{|\hat{b}|}{\hat{\sigma}}\sqrt{S_{xx}}=\frac{0.926411}{\sqrt{3.436}}\sqrt{297.6}=8.6217\geq t_{0.025}(10-2)=2.306,$$

因此接受H_1。

回归系数b的95%置信区间:

$$\left(\hat{b}-t_{0.025}(10-2)\times\frac{\hat{\sigma}}{\sqrt{S_{xx}}},\hat{b}+t_{0.025}(10-2)\times\frac{\hat{\sigma}}{\sqrt{S_{xx}}}\right)=(0.678629,\ 1.174193).$$

注意：

解题过程中可能需用到自然对数底数

$$\mathrm{e}\approx 2.7183,$$

标准正态分布函数$\Phi(\cdot)$在若干点处的函数值

$$\Phi(0.6)=0.7257,\Phi(1.65)=0.9500,\Phi(1.96)=0.9750,\Phi(2)=0.977,\Phi(3)=0.9987,$$

以及卡方χ^2分布、t-分布和F-分布的上分位数

$\chi^2_{0.025}(6)=14.440,\chi^2_{0.05}(6)=12.592,\chi^2_{0.025}(9)=19.022,\chi^2_{0.975}(9)=2.700,\chi^2_{0.05}(9)=16.919,$
$\chi^2_{0.05}(25)=37.652,\chi^2_{0.975}(24)=12.401,\chi^2_{0.975}(25)=13.120,\chi^2_{0.95}(24)=13.848,\chi^2_{0.95}(25)=14.611,$
$t_{0.025}(5)=2.571,t_{0.05}(5)=2.015,t_{0.025}(9)=2.2622,t_{0.05}(9)=1.8311,t_{0.025}(18)=2.3646,$
$t_{0.05}(18)=1.7341,t_{0.025}(8)=2.306,t_{0.05}(8)=1.8595,t_{0.025}(20)=2.0860,t_{0.025}(24)=2.0639,$
$t_{0.05}(24)=1.7109,t_{0.025}(25)=2.0595,t_{0.05}(25)=1.7081,F_{0.05}(3,6)=4.76,F_{0.05}(4,6)=4.53,$
$F_{0.05}(9,9)=3.18,F_{0.025}(9,9)=4.03,F_{0.05}(2,12)=3.89,F_{0.025}(2,12)=5.10.$

注意：$\mathrm{e}\approx 2.7183,\Phi(1.65)=0.9500,\Phi(2)=0.977,\chi^2_{0.05}(6)=12.592,\chi^2_{0.025}(9)=19.022,\chi^2_{0.95}(9)=3.325,\chi^2_{0.975}(9)=2.700,\chi^2_{0.05}(9)=16.919,F_{0.05}(9,9)=3.18,F_{0.025}(9,9)=4.03,F_{0.05}(2,12)=3.89,t_{0.025}(18)=2.3646,t_{0.025}(8)=2.306,t_{0.05}(8)=1.8595,t_{0.05}(18)=1.7341,t_{0.025}(9)=2.2622.$