2015-2016（2）概率论与数理统计-信计（B）参考答案

一、填空题〖每小题 5 分，共计 20 分〗

1、设 A,B,C 是三个随机事件，则事件 $AB\overline{C}\cup A\overline{B}C\cup\overline{A}BC$ 表达的涵义是<u>仅有两个发生或只有一个事件不发生</u>**。**

2、设随机变量 X 与 Y 相互独立，且 $X\sim U(0,4)$，$Y\sim P(3)$，则 $Var(X+3Y)=\underline{85/3}$。

3、设 $x_1,x_2\cdots,x_n$ 是来自总体 $X\sim N(\mu,\sigma^2)$ 的样本（$n>1$，σ^2 其未知），其样本均值和样本方差分别记为 $\overline{x}$ 和 s^2。则求 μ 的置信度为 $1-\alpha$ 的单侧置信区间时，所取的枢轴量为 $\underline{t=\dfrac{\overline{x}-\mu}{s/\sqrt{n}}\sim t(n-1)}$。（注：指明其服从的分布）

4、设 A,B,C 是随机试验 E 的三个事件，$P(\bullet)$ 即表示事件的概率，则

$P(A\cup B\cup C)=\underline{P(A)+P(B)+P(C)-P(AB)-P(BC)-P(AC)+P(ABC)}$ 或

$\underline{1-P(\overline{A}\overline{B}\overline{C})}$。

二、〖计 12 分〗有两箱零件：第一箱装 50 件，其中 20 件是一等品；第二箱装 30 件，其中 18 件是一等品。现从两箱中随意挑出一箱，再从该箱中先后任取两个零件，试求：
（1）第一次取出是一等品的概率；
（2）已知在第一次取出是一等品的条件下，第二次取出仍是一等品的概率。

解：记 A_i 表示取出是第 i 箱的事件，B_i 表示第 i 取出的是一等品的事件（$i=1,2$）。

由条件知：$P(A_1)=P(A_2)=\dfrac{1}{2}$。

（1）$P(B_1)=P(A_1)\cdot P(B_1\mid A_1)+P(A_2)\cdot p(B_1\mid A_2)=\dfrac{1}{2}\cdot\dfrac{20}{50}+\dfrac{1}{2}\cdot\dfrac{18}{30}=\dfrac{1}{2}$。

（2）
$$P(B_2\mid B_1)=\frac{P(B_1B_2)}{P(B_1)}=\frac{P(A_1B_1B_2)+P(A_2B_1B_2)}{P(B_1)}$$
$$=\frac{P(A_1)P(B_1\mid A_1)P(B_2\mid A_1B_1)+P(A_2)P(B_1\mid A_2)P(B_2\mid A_2B_1)}{P(B_1)}$$
$$=\frac{20}{50}\cdot\frac{19}{49}+\frac{18}{30}\cdot\frac{17}{29}=\frac{3601}{7105}。$$

三、〖计 12 分〗设 $x_1,\cdots,x_n$ 是来自总体 X 的一个样本，且 $X\sim Exp(\lambda)$。试求：$Y=\min\{x_1,\cdots,x_n\}$ 的概率密度函数 $p_Y(y)$。

解：由 X 的分布函数为 $F_X(x)=\int_{-\infty}^{x}p_X(x)dx=\begin{cases}1-e^{-\lambda x}, & x\ge 0\\ 0, & x<0\end{cases}$。

$\forall y\in R$

当 $y \le 0$ 时，$F_Y(y)=0$；

当 $y>0$ 时，$F_Y(y)=P(Y=\min\{x_1,x_2,\cdots,x_n\}\le y)$

$$=1-P(Y=\min\{x_1,x_2,\cdots,x_n\}>y)$$
$$=1-P(x_1>y,x_2>y,\cdots,x_n>y)$$
$$=1-\prod_{k=1}^{n}P(x_k>y)=1-e^{-n\lambda y}\text{；}$$

于是，$F_Y(y)=\begin{cases}1-e^{-n\lambda y}, & y\ge 0\\ 0, & y<0\end{cases}$。故有

$$P_Y(y)=\begin{cases}n\lambda e^{-n\lambda y}, & y\ge 0\\ 0, & y<0\end{cases}\text{。}$$

四、〖计 14 分〗设 $x_1,x_2,\cdots,x_n$ 是来自总体 X 的一个样本，已知随机变量 X 的期望 μ（$\mu\ne 0$）和方差 σ^2（$\sigma^2>0$）均存在。试求

（1）组合系数 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 满足什么条件时，样本 $x_1,x_2,\cdots,x_n$ 的线性组合 $\sum_{i=1}^{n}\alpha_i x_i$ 是参数 μ 的无偏估计？

（2）在（1）的基础上，当组合系数 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 取何值时，统计量 $\sum_{i=1}^{n}\alpha_i x_i$ 是 μ 的最有效估计？

解：（1）要使 $\mu=E\left(\sum_{i=1}^{n}\alpha_i x_i\right)=\sum_{i=1}^{n}E(\alpha_i x_i)=\mu\sum_{i=1}^{n}\alpha_i$，必须 $\sum_{i=1}^{n}\alpha_i=1$。因此，当组合系数 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 满足 $\sum_{i=1}^{n}\alpha_i=1$ 时，统计量 $\sum_{i=1}^{n}\alpha_i x_i$ 是 μ 的无偏估计；

（2）由 $Var\left(\sum_{i=1}^{n}\alpha_i x_i\right)=\sigma^2\sum_{i=1}^{n}\alpha_i^2$。要使在条件 $\sum_{i=1}^{n}\alpha_i=1$ 之下，函数 $Var\left(\sum_{i=1}^{n}\alpha_i x_i\right)$ 最小，即通过 Lagrange 乘数法容易求得：当 $\alpha_1=\alpha_2=\cdots=\alpha_n=\frac{1}{n}$ 时，$Var\left(\sum_{i=1}^{n}\alpha_i x_i\right)$ 达到最小，且取值为 $\frac{\sigma^2}{n}$，即此时统计量 $\sum_{i=1}^{n}\alpha_i x_i$ 是 μ 的最有效估计。

五、〖计 12 分〗已知某种合金钢中的含碳量 x（%）与强度 y（kg/cm^2）具有近似的线性关系。现测得 16 组实验数据整理如下：$\overline{x}=0.125$，$\overline{y}=45.7886$，$l_{xx}=0.3024$，$l_{xy}=25.2518$，$l_{yy}=2432.4566$。试求：（1）建立 y 关于 x 的一元线性回归方程；（2）对建立的回归方程进行显著性检验（$\alpha=0.05$）；（3）当 $x=0.15$，计算 y 的预测值。

解：（1）由$\widehat{\beta_1}=l_{xy}/l_{xx}=83.5046$，$\widehat{\beta_0}=\overline{y}-\widehat{\beta_1}\overline{x}=35.3505$，因此，$y$对$x$的线性回归方程为：$\hat{y}=35.3505+83.5046x$；

（2）由$S_R=\widehat{\beta_1}^2 l_{xx}=2108.6422$，$S_e=S_T-S_R=l_{yy}-S_R=323.8144$，可得下列方差分析表

来源	平方和	自由度	均方	F 比
回 归	$S_R=2108.6422$	$f_R=1$	$MS_R=2108.6422$	91.1664
残 差	$S_e=323.8144$	$f_e=14$	$MS_e=23.1297$	
总 计	$S_T=2432.4566$	$f_T=15$		

由于$\widehat{F}=91.1664>F_{0.95}(1,14)=4.6$，因此（1）中回归方程是显著的；

（3）当$x=0.15$，y的预测值为$\hat{y}=47.8762$。

六、〖计 12 分〗从某锌矿的东、西两支矿脉中各抽取容量分别为 9 和 8 的样本进行测试，整理可得它们的含锌量（%）的平均值及样本方差分别如下：

$$\overline{x}=23.0,\ S_x^2=1337;\ \overline{y}=26.9,\ S_y^2=1736$$

假设东、西两支矿脉的含锌量都服从正态分布，且相互独立。若取$\alpha=0.05$，试求：

（1）东支矿脉含锌量的$1-\alpha$双侧置信区间；

（2）能否认为两支矿脉含锌量的方差是相同的?

解：（1）取枢轴量为$t=\dfrac{\overline{x}-\mu_x}{s_x/\sqrt{n}}\sim t(m-1)$，于是东支矿脉含锌量的$1-\alpha$双侧置信区间为：

$$\overline{x}\pm\frac{s_x}{\sqrt{m}}t_{1-\alpha/2}(m-1)=23.0\pm\frac{\sqrt{1337}}{3}\cdot 2.3060$$
$$=23.0\pm 28.1063$$

由实际意义可知：东支矿脉含锌量的$1-\alpha$双侧置信区间为$[0,51.1063]$。

其中：$m=9$，$t_{1-\alpha/2}(m-1)=2.3060$。

（2）检验假设：$H_0:\sigma_1^2=\sigma_2^2$ vs $H_1:\sigma_1^2\neq\sigma_2^2$

取检验统计量：$F=\dfrac{s_x^2}{s_y^2}\sim F(m-1,n-1)$

其中$n=8$。

当$\alpha=0.05$时，拒绝域为

$$W=\left\{F\mid F>F_{1-\alpha/2}(m-1,n-1)\text{或}F<F_{\alpha/2}(m-1,n-1)\right\}$$
$$=\left\{F\mid F>4.9\text{或}F<1/4.53\right\}$$

其中：$F_{0.975}(8,7)=4.9$，$F_{0.025}(8,7)=1/4.53$。

由于$\widehat{F}=\dfrac{s_x^2}{s_y^2}=\dfrac{1337}{1736}=0.7702\notin W$，因此，可以认为它们的方差相等。

七、〖计 12 分〗设随机变量X的期望μ和方差σ^2（>0）均存在。$x_1,x_2,\cdots,x_n$是来自总体X的一个样本，试利用特征函数证明：

$$Y_n=\frac{\sqrt{n}\left(\bar{x}-\mu\right)}{\sigma}\xrightarrow{L}N(0,1)\text{。}$$

其中“$\xrightarrow{L}$”表示“依分布收敛”。

证明：记 $X-\mu$ 的特征函数为 $\varphi(t)$，则由特征函数性质可知：$\varphi'(0)=0$，$\varphi''(0)=-\sigma^2$。

由 $Y_n=\dfrac{\sqrt{n}\left(\bar{x}-\mu\right)}{\sigma}=\sum\limits_{i=1}^{n}\dfrac{x_i-\mu}{\sigma\sqrt{n}}$，可得 Y_n 的特征函数为

$$\varphi_{Y_n}(t)=\varphi^n\left(\frac{t}{\sigma\sqrt{n}}\right)=\left[\left(1-\frac{t^2}{2n}\right)^{-\frac{2n}{t^2}}\right]^{-\frac{t^2}{2}}$$

易得：$\lim\limits_{n\to\infty}\varphi_{Y_n}(t)=e^{-\frac{t^2}{2}}$。即当 $n\to\infty$ 时，随机变量 Y_n 依分布收敛于标准正态分布。

八、〖计 6 分〗通过本门课程的学习，从研究对象和目的来总结一下本课程。

解：本课程的研究对象为随机事件。在一次随机试验中，某随机事件发生与否不能准确预知，但在大量重复试验中，这一随机事件的发生具有统计规律。本课程的目的就是要研究它的统计规律。