2021 年随机数学与统计期末考试

mathsdream 整理版

2021.06.22

说明

题目解析见后。

点击链接可查看更多数学基础课程的近年考题：

`https://github.com/mathsdream/THU-math-source`。

一、(15 分)

设 X, Y 相互独立，均服从 $U(0,n)$，

(1) 求 $P(X+Y\le 1)$；

(2) 求 $P(\min(X,Y)\le 1)$；

(3) 令 $W = X-[x]$，其中 $[x]$ 为 x 的整数部分，求 W 的分布。

二、(15 分)

设 X 服从参数为 λ 的指数分布，$F(x)$ 为 X 的分布函数，$F(3\ln 2)=\frac{1}{2}$。

(1) 求 λ 的值；

(2) 求 $P(X-EX>\sqrt{DX})$；

(3) 随机变量 $X_1, X_2, \cdots, X_n$ 相互独立，且与 X 同分布。令 $S_n=\sum\limits_{k=1}^{n} kX_k$，证明：$\dfrac{S_n}{n(n+1)} \xrightarrow{P} \dfrac{3}{2}$。

三、(10 分)

设 $\{B_t : t\ge 0\}\,(B_0=0)$ 为标准布朗运动。

(1) 记 $U=2B_1+B_2-B_3$，求 U 的分布和 $P(|U|\le DU)$；

(2) 求 $P(B_3\le 3|B_1=1)$ 和 $P(B_1\le 1|B_3=3)$。

四、(15 分)

已知随机向量 (X,Y) 的联合分布密度为

$$f_{X,Y}(x,y)=\begin{cases}6y, & 0<y<x<1\\ 0, & \text{其它}\end{cases}$$

(1) 求 $E(Y|X)$；

(2) 求 $X-Y$ 的分布；

(3) 求 $E(X|X+Y<1)$。

五、(20 分)

设 $X_1,\cdots,X_n$ 是正态总体 $X\sim N(0,\sigma^2)$ 的一个样本，$\bar{X}$ 为其样本均值，

(1) 求 $\bar{X}$ 的矩母函数 $M_{\bar{X}}(u)$ 和特征函数 $\varphi_{\bar{X}}(\theta)$；

(2) 求 $E(2X_1+3X_2|\bar{X})$；

(3) 证明：$\dfrac{X_1}{\sqrt{\frac{1}{6}(X_2+X_3+X_4)^2+\frac{1}{4}(X_2-X_3)^2}}\sim t(2)$；

(4) 统计量 $T=\dfrac{\sqrt{\frac{\pi}{2}}}{n}\sum_{i=1}^{n}|X_i|$ 是否为 σ 的无偏估计？证明：证明：$\sqrt{n}\,(\ln T-\ln\sigma)\xrightarrow{D}N(0,\dfrac{\pi-2}{2})$。

六、(25 分)

设 $X_1,\cdots,X_n$ 是总体 X 的一个样本，X 的密度函数为：

$$f(x;\theta)=\begin{cases}\dfrac{1}{\theta}x^{\frac{1}{\theta}-1}, & 0<x<1\\ 0, & \text{其它}\end{cases}$$

其中 $\theta>0$ 为未知参数。

(1) 设统计量 $T=-\sum_{i=1}^{n}\ln X_i$，试问 T 是否为充分统计量？是否为完备统计量？为什么？

(2) 证明：$T\sim \mathrm{Gamma}(n,\dfrac{1}{\theta})$；

(3) 求参数 θ 的最大似然估计 $\hat{\theta}_{MLE}$，其是否为 θ 的 UMVUE？

(4) 构造枢轴量，并求出参数 θ 的置信水平为 $1-\alpha$ 的等尾置信区间；

(5) 考虑假设检验问题：$H_0:\theta=1$，$H_1:\theta=2$，进行似然比检验，给出拒绝域的表达式。

解析为个人做本卷时的第一思路，不代表最巧妙的解法，同时不保证完全无误，仅供参考。

一、

设 X,Y 相互独立，均服从 $U(0,n)$，

(1) 求 $P(X+Y\leq 1)$；

(2) 求 $P(\min(X,Y)\leq 1)$；

(3) 令 $W=X-[x]$，其中 $[x]$ 为 x 的整数部分，求 W 的分布。

解析：

(1) X,Y 的联合密度函数为 $f_{X,Y}(x,y)=f_X(x)f_Y(y)=\dfrac{1}{n^2}$，$0<x,y<n$，所以

$$P(X+Y\leq 1)=\int_0^1\int_0^{1-x}\frac{1}{n^2}dydx=\frac{1}{2n^2}$$

(2)

$$\begin{aligned}P(\min(X,Y)\leq 1)&=1-P(\min(X,Y)>1)\\&=1-P(X>1,Y>1)\\&=1-P(X>1)P(Y>1)\\&=1-(1-\frac{1}{n})(1-\frac{1}{n})\\&=\frac{2}{n}-\frac{1}{n^2}\end{aligned}$$

(3) 先求 W 的分布函数：

$$\begin{aligned}F_W(w)=P(W\leq w)&=P(X-[X]\leq w)\\&=\sum_{k=0}^{n-1}P(k\leq X\leq k+w)\\&=\sum_{k=0}^{n-1}\int_k^{k+w}\frac{1}{n}dx\\&=\sum_{k=0}^{n-1}\frac{w}{n}=\frac{nw}{n}=w\quad(0<w<1)\end{aligned}$$

所以 W 的密度函数为 $f_W(w)=1$，$0<w<1$，即 $W\sim U(0,1)$。

二、

设 X 服从参数为 λ 的指数分布，$F(x)$ 为 X 的分布函数，$F(3\ln 2)=\dfrac{1}{2}$。

(1) 求 λ 的值；

(2) 求 $P(X-EX>\sqrt{DX})$；

(3) 随机变量 $X_1,X_2,\cdots,X_n$ 相互独立，且与 X 同分布。令 $S_n=\sum\limits_{k=1}^{n}kX_k$，证明：$\dfrac{S_n}{n(n+1)}\xrightarrow{P}\dfrac{3}{2}$。

解析：

(1) $F(x)=1-e^{-\lambda x}$，所以 $F(3\ln 2)=1-e^{-3\lambda\ln 2}=\frac{1}{2}$，因此 $\lambda=\frac{1}{3}$。

(2)

$$
\begin{aligned}
EX=\frac{1}{\lambda}=3,\quad DX=\frac{1}{\lambda^2}=9\\
P(X-EX>\sqrt{DX})&=P(X>6)\\
&=\int_6^{\infty}\frac{1}{3}e^{-\frac{1}{3}x}dx\\
&=e^{-2}
\end{aligned}
$$

(3)

$$
\begin{aligned}
&ES_n=E\left(\sum_{k=1}^{n}kX_k\right)=\sum_{k=1}^{n}kE(X_k)=\sum_{k=1}^{n}k\cdot 3=3\cdot\frac{n(n+1)}{2}\\
&DS_n=D\left(\sum_{k=1}^{n}kX_k\right)=\sum_{k=1}^{n}k^2D(X_k)=\sum_{k=1}^{n}k^2\cdot 9=3\cdot\frac{n(n+1)(2n+1)}{2}\\
&E\left(\frac{S_n}{n(n+1)}\right)=\frac{3}{2}\\
&D\left(\frac{S_n}{n(n+1)}\right)=\frac{1}{(n(n+1))^2}\frac{3n(n+1)(2n+1)}{2}=\frac{3(2n+1)}{2n(n+1)}
\end{aligned}
$$

由切比雪夫不等式：

$$
P\left(\left|\frac{S_n}{n(n+1)}-\frac{3}{2}\right|>\epsilon\right)\le\frac{3(2n+1)}{2n(n+1)\epsilon^2}\xrightarrow{n\to\infty}0
$$

所以 $\frac{S_n}{n(n+1)}\xrightarrow{P}\frac{3}{2}$。

三、

设 $\{B_t:t\ge 0\}\,(B_0=0)$ 为标准布朗运动。

(1) 记 $U=2B_1+B_2-B_3$，求 U 的分布和 $P(|U|\le DU)$；

(2) 求 $P(B_3\le 3|B_1=1)$ 和 $P(B_1\le 1|B_3=3)$。

解析：

(1) $U=2B_1+B_2-B_3=(2,1,-1)\,(B_1,B_2,B_3)^T$，而

$$
(B_1,B_2,B_3)^T\sim N(\boldsymbol{\mu},\boldsymbol{\Sigma})=N\left(\mathbf{0},\begin{pmatrix}1&1&1\\1&2&2\\1&2&3\end{pmatrix}\right)
$$

所以 $U\sim N(0,(2,1,-1)\,\boldsymbol{\Sigma}\,(2,1,-1)^T)=N(0,5)$。

$$
\begin{aligned}
P(|U|\le DU)&=P(-DU\le U\le DU)\\
&=P\left(-\sqrt{5}\le\frac{U}{\sqrt{5}}\le\sqrt{5}\right)\\
&=\Phi(\sqrt{5})-\Phi(-\sqrt{5})\\
&=2\Phi(\sqrt{5})-1
\end{aligned}
$$

(2) 由独立增量性质，$B_3 - B_1$ 与 B_1 独立，且 $B_3 - B_1 \sim N(0,2)$，所以

$$\begin{aligned}P(B_3 \le 3|B_1 = 1) &= P(B_3 - B_1 \le 2|B_1 = 1)\\&= P(B_3 - B_1 \le 2)\\&= P\left(\frac{B_3 - B_1}{\sqrt{2}} \le \frac{2}{\sqrt{2}}\right)\\&= \Phi(\sqrt{2})\end{aligned}$$

此外，由 $(B_1, B_3) \sim N(0, 0, \frac{1}{\sqrt{3}}, 1, 3)$，所以 $B_1|B_3 = 3 \sim N(1, \frac{2}{3})$，因此

$$\begin{aligned}P(B_1 \le 1|B_3 = 3) &= P\left(\frac{B_1 - 1}{\sqrt{\frac{2}{3}}} \le \frac{0}{\sqrt{\frac{2}{3}}}|B_3 = 3\right)\\&= \Phi(0) = \frac{1}{2}\end{aligned}$$

四、

设 (X,Y) 的联合分布密度为

$$f_{X,Y}(x,y) = \begin{cases}6y, & 0 < y < x < 1\\0, & 其它\end{cases}$$

(1) 求 $E(Y|X)$；

(2) 求 $X - Y$ 的分布；

(3) 求 $E(X|X + Y < 1)$。

解析：

(1)

$$f_X(x) = \int_0^x f_{X,Y}(x,y)dy = \int_0^x 6ydy = 3x^2 \quad (0 < x < 1)$$

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{2y}{x^2} \quad (0 < y < x < 1)$$

$$E(Y|X = x) = \int_0^x yf_{Y|X}(y|x)dy = \int_0^x y \cdot \frac{2y}{x^2}dy = \frac{2x}{3}$$

所以 $E(Y|X) = \frac{2X}{3}$。(2) 记 $U = X - Y, V = X$，则 $X = V, Y = V - U$，所以

$$\begin{aligned}f_{U,V}(u,v) &= f_{X,Y}(v, v-u)\left|\frac{\partial(X,Y)}{\partial(U,V)}\right|\\&= 6(v-u) \quad (0 < u < v < 1)\end{aligned}$$

所以

$$\begin{aligned}f_U(u) &= \int_u^1 f_{U,V}(u,v)dv = \int_u^1 6(v-u)dv\\&= 3(1-u)^2 \quad (0 < u < 1)\end{aligned}$$

(3)

$$\begin{aligned}E(X|X+Y<1) &= \frac{E(X\cdot I_{X+Y<1})}{P(X+Y<1)}\\ &= \frac{\int_0^{\frac{1}{2}}\int_y^{1-y} x\cdot 6ydxdy}{\int_0^{\frac{1}{2}}\int_y^{1-y} 6ydxdy}\\ &= \frac{1}{2}\end{aligned}$$

五、

设 $X_1,\cdots,X_n$ 是正态总体 $X\sim N(0,\sigma^2)$ 的一个样本，$\bar{X}$ 为其样本均值，

(1) 求 $\bar{X}$ 的矩母函数 $M_{\bar{X}}(u)$ 和特征函数 $\varphi_{\bar{X}}(\theta)$；

(2) 求 $E(2X_1+3X_2|\bar{X})$；

(3) 证明：$\dfrac{X_1}{\sqrt{\frac{1}{6}(X_2+X_3+X_4)^2+\frac{1}{4}(X_2-X_3)^2}}\sim t(2)$；

(4) 统计量 $T=\dfrac{\sqrt{\frac{\pi}{2}}}{n}\sum_{i=1}^{n}|X_i|$ 是否为 σ 的无偏估计？证明：$\sqrt{n}\,(\ln T-\ln\sigma)\xrightarrow{D}N(0,\dfrac{\pi-2}{2})$。

解析：

(1) $\bar{x}\sim N(0,\dfrac{\sigma^2}{n})$，所以

$$M_{\bar{X}}(u)=\exp\left(\frac{\sigma^2u^2}{2n}\right),\quad \varphi_{\bar{X}}(\theta)=\exp\left(-\frac{\sigma^2\theta^2}{2n}\right)$$

(2) 由 $(X_1,\cdots,X_n)^T$ 服从高斯分布和高斯分布的线性性，可以得到 $(2X_1+3X_2,\bar{X})$ 服从二元正态分布，参数为

$$\begin{aligned}&E(2X_1+3X_2)=0,\quad E(\bar{X})=0\\ &D(2X_1+3X_2)=13\sigma^2,\quad D(\bar{X})=\frac{\sigma^2}{n}\\ &r(2X_1+3X_2,\bar{X})=\frac{cov(2X_1+3X_2,\bar{X})}{\sqrt{D(2X_1+3X_2)D(\bar{X})}}=\frac{5\sqrt{n}}{\sqrt{13}}\end{aligned}$$

从而得到 $E(2X_1+3X_2|\bar{X})=5\bar{X}$。

(3) 记

$$\begin{aligned}Y_1&=\frac{X_1}{\sigma}\sim N(0,1)\\ Y_2&=\frac{X_2+X_3+X_4}{\sqrt{3}\sigma}\sim N(0,1)\\ Y_3&=\frac{X_2-X_3}{\sqrt{2}\sigma}\sim N(0,1)\end{aligned}$$

同时，显然 (Y_2,Y_3) 呈高斯分布，而 $Cov(Y_2,Y_3)=0$，所以 (Y_2,Y_3) 独立。所以

$$Z=Y_2^2+Y_3^2\sim\chi^2(2)$$

而 Y_1 与 Z 独立，所以

$$\frac{X_1}{\sqrt{\frac{1}{6}(X_2+X_3+X_4)^2+\frac{1}{4}(X_2-X_3)^2}}=\frac{Y_1}{\sqrt{Z/2}}\sim t(2)$$

(4)

$$E(|X_i|) = \sqrt{\frac{2}{\pi}}\sigma$$
$$D(|X_i|) = \left(1 - \frac{2}{\pi}\right)\sigma^2$$

所以 $E(T) = \frac{\sqrt{\frac{\pi}{2}}}{n} \cdot n \cdot \sqrt{\frac{2}{\pi}}\sigma = \sigma$，因此 T 是 σ 的无偏估计。

由中心极限定理，$\sqrt{n}(T - \sigma) \xrightarrow{D} N(0, \frac{\pi-2}{2}\sigma^2)$，由 Delta 方法，取 $g(x) = \ln x$，$g'(\sigma) = \frac{1}{\sigma}$，得到

$$\sqrt{n}(\ln T - \ln\sigma) \xrightarrow{D} N(0, g'(\sigma)^2\frac{\pi-2}{2}\sigma^2) = N(0, \frac{\pi-2}{2})$$

六、

设 $X_1, \cdots, X_n$ 是总体 X 的一个样本，X 的密度函数为：

$$f(x;\theta) = \begin{cases} \frac{1}{\theta}x^{\frac{1}{\theta}-1}, & 0 < x < 1 \\ 0, & \text{其它} \end{cases}$$

其中 $\theta > 0$ 为未知参数。

(1) 设统计量 $T = -\sum_{i=1}^{n} \ln X_i$，试问 T 是否为充分统计量？是否为完备统计量？为什么？

(2) 证明：$T \sim \text{Gamma}(n, \frac{1}{\theta})$；

(3) 求参数 θ 的最大似然估计 $\hat{\theta}_{MLE}$，其是否为 θ 的 UMVUE？

(4) 构造枢轴量，并求出参数 θ 的置信水平为 $1-\alpha$ 的等尾置信区间；

(5) 考虑假设检验问题：$H_0 : \theta = 1$，$H_1 : \theta = 2$，进行似然比检验，给出拒绝域的表达式。

解析：

(1)

$$\begin{aligned} f(x_1, \cdots, x_n; \theta) &= \prod_{i=1}^{n} f(x_i; \theta) \\ &= \prod_{i=1}^{n} \frac{1}{\theta} x_i^{\frac{1}{\theta}-1} I(0 < x_i < 1) \\ &= \frac{1}{\theta^n} (\prod_{i=1}^{n} x_i)^{\frac{1}{\theta}-1} \prod_{i=1}^{n} I(0 < x_i < 1) \\ &= \frac{1}{\theta^n} e^{-T(\frac{1}{\theta}-1)} \prod_{i=1}^{n} I(0 < x_i < 1) \end{aligned}$$

这是一个指数族，由此可知，T 是 θ 的充分完备统计量。

(2) 设 $Y_i = -\ln X_i$，则

$$f_{Y_i}(y_i) = f_{X_i}(e^{-y_i}) \left| \frac{\partial X_i}{\partial Y_i} \right| = \frac{1}{\theta} e^{-\frac{y_i}{\theta}}, \quad y_i > 0$$

所以 $Y_i \sim \mathrm{Exp}(\frac{1}{\theta})$，$Y_1, \cdots, Y_n$ 相互独立，所以

$$T = \sum_{i=1}^{n} Y_i \sim \mathrm{Gamma}(n, \frac{1}{\theta})$$

(3)

$$\ln L(\theta) = -n \ln \theta - T(\frac{1}{\theta} - 1)$$
$$\frac{\partial \ln L(\theta)}{\partial \theta} = 0 \Rightarrow \hat{\theta}_{MLE} = \frac{T}{n}$$

已知 T 是 θ 的充分完备统计量。$E(\hat{\theta}_{MLE}) = \theta$，所以 $\hat{\theta}_{MLE}$ 是 θ 的无偏估计，同时 $\hat{\theta}_{MLE}$ 是 T 的函数，所以 $\hat{\theta}_{MLE}$ 是 θ 的 UMVUE。

(4) 由 $T \sim \mathrm{Gamma}(n, \dfrac{1}{\theta})$，所以 $2T/\theta \sim \chi^2(2n)$，因此

$$P\left(\chi^2_{\frac{\alpha}{2}}(2n) < \frac{2T}{\theta} < \chi^2_{1-\frac{\alpha}{2}}(2n)\right) = 1 - \alpha$$
$$P\left(\frac{2T}{\chi^2_{1-\frac{\alpha}{2}}(2n)} < \theta < \frac{2T}{\chi^2_{\frac{\alpha}{2}}(2n)}\right) = 1 - \alpha$$

所以置信区间为 $\left(\dfrac{2T}{\chi^2_{1-\frac{\alpha}{2}}(2n)}, \dfrac{2T}{\chi^2_{\frac{\alpha}{2}}(2n)}\right)$。

(5) 似然比 $\Lambda = \dfrac{L(2)}{L(1)} = 2^{-n} e^{\frac{T}{2}}$，拒绝域为 $W = \{\Lambda \geq c\} = \{2T \geq d\}$，其中 c, d 是常数，取显著性水平为 α，则

$$P\left(2T \geq d | \theta = 1\right) = \alpha$$

由 $2T|\theta = 1 \sim \chi^2(2n)$，所以拒绝域为 $\left\{2T \geq \chi^2_{1-\alpha}(2n)\right\} = \left\{T \geq \dfrac{1}{2}\chi^2_{1-\alpha}(2n)\right\}$。