概率论第 5 次习题课讲义

宗语轩

2022 秋, USTC

1 作业讲评

4.2.4 由题意得,

$$\mathbb{P}(Y(y)=k)=\mathbb{P}(X_1\leqslant y\cdots,X_{k-1}\leqslant y,X_k>y)=\prod_{i=1}^{k-1}\mathbb{P}(X_i\leqslant y)\mathbb{P}(X_k>y)=(F(y))^{k-1}(1-F(y))$$

因此

$$\mathbb{P}(Y(y)\geqslant k)=(F(y))^{k-1}\quad\Longrightarrow\quad \mathbb{E}[Y(y)]=\sum_{k=1}^{\infty}\mathbb{P}(Y(y)\geqslant k)=\sum_{k=0}^{\infty}(F(y))^k=\frac{1}{1-F(y)}.$$

故

$$\mathbb{P}(Y(y)>\mathbb{E}[Y(y)])=\mathbb{P}\left(Y(y)>\frac{1}{1-F(y)}\right)=\mathbb{P}\left(Y(y)\geqslant\left[\frac{1}{1-F(y)}\right]+1\right)=(F(y))^{[\frac{1}{1-F(y)}]+1}.$$

利用

$$\frac{1}{1-F(y)}\leqslant\left[\frac{1}{1-F(y)}\right]+1<\frac{1}{1-F(y)}+1,$$

及 $\lim\limits_{y\to+\infty}F(y)=1$, 有

$$(F(y))^{\frac{1}{1-F(y)}+1}<\mathbb{P}(Y(y)>\mathbb{E}[Y(y)])\leqslant(F(y))^{\frac{1}{1-F(y)}}.$$

令 $y\to+\infty$, 两边极限均为 e^{-1}, 夹逼即得

$$\lim_{y\to+\infty}\mathbb{P}(Y(y)>\mathbb{E}[Y(y)])=\mathrm{e}^{-1}\quad\Longrightarrow\quad\lim_{y\to+\infty}\mathbb{P}(Y(y)\leqslant\mathbb{E}[Y(y)])=1-\mathrm{e}^{-1}.$$

注. $\mathbb{E}[Y(y)]$ 不是整数, 需要做进一步处理.

4.5.7

$$\begin{aligned}Cov(\overline{X},X_r-\overline{X})&=Cov\left(\frac{1}{n}\sum_{k=1}^{n}X_k,X_r-\frac{1}{n}\sum_{k=1}^{n}X_k\right)=Cov\left(\frac{1}{n}\sum_{k=1}^{n}X_k,-\frac{1}{n}\sum_{k=1,k\neq r}^{n}X_k+\left(1-\frac{1}{n}\right)X_r\right)\\&=-\frac{1}{n^2}\sum_{k=1,k\neq r}^{n}Var(X_k)+\frac{1}{n}\left(1-\frac{1}{n}\right)Var(X_r)=0.\end{aligned}$$

[0]个人主页: http://home.ustc.edu.cn/~zyx240014/.
发现错误欢迎联系: zyx240014@mail.ustc.edu.cn.

另解: 由对称性知

$$d := Cov(\overline{X}, X_1 - \overline{X}) = Cov(\overline{X}, X_2 - \overline{X}) = \cdots = Cov(\overline{X}, X_n - \overline{X})$$

因此

$$nd = \sum_{k=1}^{n} Cov(\overline{X}, X_k - \overline{X}) = Cov(\overline{X}, \sum_{k=1}^{n}(X_k - \overline{X})) = Cov(\overline{X}, 0) = 0 \implies d = 0.$$

注. 注意区分相同指标和不同指标的协方差计算.

4.6.6 注:

$$\mathbb{P}(N > n) = \mathbb{P}(X_1 = X_2 + \cdots + X_n \leqslant x).$$

因为 X_i 是 i.i.d 且 $U(0,1)$ 的, 因此该模型等同于求 n 维单形的测度, 归纳即可.

2 专题选讲

2.1 期望浅谈

在概率论中, 对分布函数的积分 $\int f dF$ 都是指 Lebesgue 积分, 而不是 Riemann 积分. 这两者有根本区别, 后者是对函数的定义域 "分蛋糕", 而前者是对函数的值域 "分蛋糕"(回顾一下可测函数 (随机变量) 的内容). 同时注意到, 可测函数对函数的连续性根本不作要求. 从动机讲, 为了把积分对象扩充到更大的一类函数——可测函数类上, 而 Riemann 积分就不够用了, 因此必须要在此基础上进一步推广, 同时注意到可测函数在极限运算下封闭, 为保持可测函数的结构和运算, 于是就据此建立了 Lebesgue 积分这一套理论. 具体内容可以参考实分析教材.

我们回顾一下课堂中已讲过的内容: 定义一般随机变量 X 的期望 $\mathbb{E}[X]$ 的过程——三步走战略 (若第一步是示性函数则为四步走).

$$\text{非负简单随机变量} \quad \longrightarrow \quad \text{非负随机变量} \quad \longrightarrow \quad \text{一般随机变量}$$

$$X = \sum_{i=1}^{n} x_i I_{A_i}, \ X_n = nI_{A_n} + \sum_{j=1}^{n\cdot 2^n} \frac{j-1}{2^n} I_{A_{n,j}} \xrightarrow{n\to+\infty} X, \ X = X^+ - X^-.$$

其中对 $X = \sum_{i=1}^{n} x_i I_{A_i}$,

$$\mathbb{E}[X] = \sum_{i=1}^{n} x_i \mathbb{P}(A_i).$$

X 可积指期望 $\mathbb{E}[|X|]$ 存在 $\iff \mathbb{E}[X^+], \mathbb{E}[X^-]$ 均存在.

以上过程可以借助下面几个 **Facts** 来理解:

1. 随机变量 $X : \Omega \to \mathbb{R}$ 是实值 Borel 可测函数: $\forall x \in \mathbb{R}, \{X \leqslant x\} \in \mathcal{F}$. **动机**: 对于随机变量, 相比过程 (Ω), 更在乎输出结果.

2. 随机变量在极限运算下封闭: $\forall n \in \mathbb{N}^*$, X_n 是随机变量且 $X := \lim_{n\to+\infty} X_n$ 存在, 则 X 也是随机变量 ($X = \liminf_{n\to\infty} X_n = \limsup_{n\to\infty} X_n$ 是随机变量).

3. X 未必是连续型的.

期望的一些**性质**: 设 X, Y, X_n 是随机变量, a, b, c 是常数, 则有

(1) 非负性: 当 $X \geqslant 0$ 时, $\mathbb{E}[X] \geqslant 0$.

(2) 归一性: $\mathbb{E}[c] = c, c \in \mathbb{R}$.

(3) 线性性: $\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$.

(4) 单调性: 当 $X \leqslant Y$ 时, $\mathbb{E}[X] \leqslant \mathbb{E}[Y]$.

(5) 绝对值不等式: $|\mathbb{E}[X]| \leqslant \mathbb{E}[|X|]$.

(6) 单调收敛定理: $X_{n+1}(\omega) \geqslant X_n(\omega) \geqslant 0, \forall n \in \mathbb{N}^*$, 且 $X_n(\omega) \to X(\omega), n \to +\infty, \omega \in \Omega$ 则

$$\mathbb{E}[X_n] \to \mathbb{E}[X], n \to \infty.$$

这里 $\mathbb{E}[X] \leqslant +\infty$.

(7) 控制收敛定理: $X_n \to X, |X_n| \leqslant Y, \forall n \in \mathbb{N}^*$, 当 $\mathbb{E}[Y] < \infty$ 时

$$\mathbb{E}[X_n] \to \mathbb{E}[X], n \to \infty,$$

其中 Y 为常数 c 时又称有界收敛定理.

(8) **Fatou** 引理: $X_n \geqslant 0$, 则

$$\mathbb{E}[\liminf_{n\to\infty} X_n] \leqslant \liminf_{n\to\infty} \mathbb{E}[X_n].$$

后三个性质证明参考群文件讲义《“再谈期望” 补充》

注 1. 由上述期望定义知, 若 $\mathbb{P}(A) = 0, \forall A \in \mathcal{F}$, 则有 $\displaystyle\int_A X \mathrm{d}F = 0$.

注 2. 由**注 1** 可知, 在原集合基础上挖去概率测度为 0 的集合并不会改变积分值, 因此很多性质我们只需考虑几乎处处 (a.s.) 条件下满足即可: $\exists \Omega_0, \mathbb{P}(\Omega_0) = 0, \forall \omega \in \Omega \backslash \Omega_0$

定理 2.1 (逐项积分定理)**.** 设随机变量 $X_n \geqslant 0$, 则有

$$\mathbb{E}[\sum_{k=1}^{\infty} X_k] = \sum_{k=1}^{\infty} \mathbb{E}[X_k].$$

提示. $S_n = \displaystyle\sum_{k=1}^{n} X_k \uparrow \sum_{k=1}^{\infty} X_k$, 再利用单调收敛定理即可.

在第 3 次习题课讲义里已经证明 (用**定理 2.1** 亦可): 对于非负整值随机变量 X, 有 $\mathbb{E}[X] = \displaystyle\sum_{n=1}^{\infty} \mathbb{P}(X \geqslant n)$. 课堂上我们已经证明: 对于连续型随机变量 X, 我们亦有

$$\mathbb{E}[X] = \int_0^{\infty} (1 - F(x))\mathrm{d}x - \int_{-\infty}^{0} F(x)\mathrm{d}x.$$

其中 F 是 X 的分布函数. 对于一般随机变量, 我们可以进行估计:

引理 2.1. 设 X 是随机变量, 则有

$$\sum_{n=1}^{\infty}\mathbb{P}(|X|\geqslant n)\leqslant\mathbb{E}(|X|)\leqslant\sum_{n=1}^{\infty}\mathbb{P}(|X|\geqslant n)+1$$

证明. 注意到

$$\sum_{n=1}^{\infty}\mathbb{P}(|X|\geqslant n)=\sum_{n=1}^{\infty}\mathbb{P}([|X|]\geqslant n)=\mathbb{E}([|X|])\leqslant\mathbb{E}(|X|)\leqslant\mathbb{E}([|X|]+1)=\sum_{n=1}^{\infty}\mathbb{P}(|X|\geqslant n)+1.$$

即可. □

例 2.1. 设 $X\geqslant 0$, 证明

$$\mathbb{E}[X]=0\iff X\xlongequal{a.s.}0.$$

证明. $\Longleftarrow$: 注意到上述**注 1** 即可;

$\Longrightarrow$: 即证 $\mathbb{P}(X>0)=0$. 而 $\{X>0\}=\bigcup\limits_{n=1}^{\infty}\left\{X>\dfrac{1}{n}\right\}$. 故有

$$\mathbb{P}\left(X>\frac{1}{n}\right)=\mathbb{E}[I_{\{X>\frac{1}{n}\}}]=n\mathbb{E}\left[\frac{1}{n}I_{\{X>\frac{1}{n}\}}\right]\leqslant n\mathbb{E}[XI_{\{X>\frac{1}{n}\}}]\leqslant n\mathbb{E}[X]=0.$$

由概率测度的次可加性知,

$$\mathbb{P}(X>0)\leqslant\sum_{n=1}^{\infty}\mathbb{P}\left(X>\frac{1}{n}\right)=0.$$

□

注. 类似上述方法可以证明: 若 X 在 Ω 上非负可积, 则 X 在 Ω 上 $a.s.$ 有限. (注意到 $\{X=+\infty\}=\bigcap\limits_{k=1}^{\infty}\{X>k\}$ 即可)

例 2.2. 设随机变量 X 可积, $\mathbb{P}(A_n)\to 0$, 证明: $\displaystyle\int_{A_n}X\mathrm{d}F\to 0$.

证明. 只需证: $\displaystyle\int_{A_n}|X|\mathrm{d}F\to 0$. 做截断: 定义随机变量

$$X_k=|X|I_{\{|X|<k\}}+kI_{\{|X|\geqslant k\}}.$$

利用单调收敛定理, 有

$$0\leqslant X_k\uparrow|X|\implies\mathbb{E}[X_k]\uparrow\mathbb{E}[|X|].$$

因此

$$\int_{A_n}|X|\mathrm{d}F=\int_{A_n}X_k\mathrm{d}F+\int_{A_n}(|X|-X_k)\mathrm{d}F\leqslant k\mathbb{P}(A_n)+\mathbb{E}[|X|-X_k].$$

对 $\forall\varepsilon>0$, 取 $k=k_0$, 使得 $\mathbb{E}[|X|-X_{k_0}]<\dfrac{\varepsilon}{2}$. 且 $\exists N\in\mathbb{N}^*$, 当 $n>N$ 时, 有

$$\mathbb{P}(A_n)<\frac{\varepsilon}{2k_0}.$$

回到上式, 则有

$$\int_{A_n}|X|\mathrm{d}F<k_0\cdot\frac{\varepsilon}{2k_0}+\frac{\varepsilon}{2}=\varepsilon.$$

□

注 1. 体会一下这题用的截断法, 在后续讲到随机变量的收敛中这是一个常用的 technique.

注 2. 本题不能直接用控制收敛定理推得, 因为 XI_{A_n} 不 $a.s.$ 收敛于 0 (反例请读者自行思考).

2.2 连续型随机变量

随机向量间的变换: 以二元为例, $(X_1,X_2)\to(Y_1,Y_2)=T(X_1,X_2)$, $y_i=y_i(x_1,x_2), i=1,2$, $T:D\subset\mathbb{R}^2\to R\subset\mathbb{R}^2, 1-1$ 映射, $(x_1,x_2)\to(y_1,y_2)$, 其逆 $x_i=x_i(y_1,y_2)$ 有连续偏导数.

定理 2.2. (X_1,X_2) 有密度 $f(x_1,x_2)$, 则

$$f_{Y_1,Y_2}(y_1,y_2)=f(x_1(y_1,y_2),x_2(y_1,y_2))\,|J(y_1,y_2)|\,I_R$$

这里 $J=\begin{vmatrix}\dfrac{\partial x_1}{\partial y_1} & \dfrac{\partial x_2}{\partial y_1}\\ \dfrac{\partial x_1}{\partial y_2} & \dfrac{\partial x_2}{\partial y_2}\end{vmatrix}$

证明. 本质变量替换, 设 $A\subset D, B=T(A)\subset T(D)=R$, 则 $(X_1,X_2)\in A\Leftrightarrow(Y_1,Y_2)\in B$, 则

$$\begin{aligned}\mathbb{P}((Y_1,Y_2)\in B)&=\mathbb{P}((X_1,X_2)\in A)\\&=\iint_A f(x_1,x_2)\mathrm{d}x_1\mathrm{d}x_2\\&=\iint_B f(x_1(y_1,y_2),x_2(y_1,y_2))\,|J|\,\mathrm{d}y_1\mathrm{d}y_2\end{aligned}$$

取 $B=T(D)\cap(-\infty,y_1]\times(-\infty,y_2]$ □

注 1. 若 $D_0\subset D,\mathbb{P}((X_1,X_2)\in D_0)=1,T$ 在 D_0 上单射, 结论亦对.

注 2. 这里二元情形可以一般化地推广到 n 元情形.

例 2.3. 设 X,Y 独立同 $N(0,1)$, 令

$$X=R\cos\Theta, Y=R\sin\Theta, R\geq 0, 0\leq\Theta\leq 2\pi$$

求 (R,Θ) 的联合密度.

解. $f(x,y)=\dfrac{1}{2\pi}e^{-\frac{1}{2}(x^2+y^2)}$, $|J|=r$, 则

$$f_{R,\Theta}(r,\theta)=\frac{1}{2\pi}re^{-\frac{1}{2}r^2}, r>0,\theta\in[0,2\pi)$$

表明 R 与 Θ 独立, 且 $\Theta\sim U[0,2\pi), f_R(r)=re^{-\frac{1}{2}r^2}, r>0$ □

副产品: 产生独立正态随机数, 设 U_1,U_2 独立同 $U[0,1]$, 令

$$X=\sqrt{-2\log U_1}\cos(2\pi U_2)$$

$$Y=\sqrt{-2\log U_1}\sin(2\pi U_2)$$

则 X 与 Y 独立同 $N(0,1)$.

例 2.4. X,Y 独立同 $N(0,1)$, 令 $U=X, V=\rho X+\sqrt{1-\rho^2}Y, \rho\in(-1,1)$, 则 (U,V) 服从二元标准正态分布. 即

$$f_{U,V}(u,v)=\frac{1}{2\pi\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2(1-\rho^2)}(x^2-2\rho xy+y^2)\right\}.$$

例 2.5 (次序统计量)**.** 随机变量 $X_1,\cdots,X_n$ i.i.d $\sim F(x)$, 其密度函数为 $f(x)$. 将 $X_1,\cdots,X_n$ 从小到大排列为 $X_{(1)}\leqslant X_{(2)}\leqslant\cdots\leqslant X_{(n)}$.

(1) 证明:
$$\mathbb{P}(X_{(n)}<x)=(F(x))^n,\quad \mathbb{P}(X_{(1)}>x)=(1-F(x))^n.$$
由此可以分别求出 $X_{(1)},X_{(n)}$ 的密度函数.

(2) 证明: n 个次序统计量 $(X_{(1)},X_{(2)},\cdots,X_{(n)})$ 的联合分布
$$\mathbb{P}(X_{(1)}<x_1,\cdots,X_{(n)}<x_n)=\begin{cases} n!\mathbb{P}(X_1<x_1,\cdots X_n<x_n,X_1<X_2<\cdots<X_n), & x_1<\cdots<x_n,\\ 0, & \text{else}.\end{cases}$$
由此导出其联合密度
$$g(x_1,\cdots,x_n)=\begin{cases} n!f(x_1)\cdots f(x_n), & x_1<\cdots<x_n,\\ 0, & \text{else}.\end{cases}$$
注. 这里 $X_{(i)}=x_i$ 已给定次序, 故 *else* 情况下为 0.

(3) 对 $\forall 1\leqslant i<j\leqslant n$, 求单个统计量 $X_{(i)}$ 的密度函数 $g_i(x)$ 和两个次序统计量 $(X_{(i)},X_{(j)})$ 的联合密度函数 $g_{ij}(x,y)$ (仅考虑 $x<y$).

(4) (留作习题) 若 X_i i.i.d. $\sim U(0,1)$, 定义极差: $R=X_{(n)}-X_{(1)}$, 求 R 的密度函数.

解. (1)
$$\mathbb{P}(X_{(n)}<x)=\mathbb{P}(X_1<x,\cdots,X_n<x)=\prod_{k=1}^n\mathbb{P}(X_k<x)=(F(x))^n,$$
$$\mathbb{P}(X_{(1)}>x)=\mathbb{P}(X_1>x,\cdots,X_n>x)=\prod_{k=1}^n\mathbb{P}(X_k>x)=(1-F(x))^n.$$
(2) 仅考虑 $x_1<\cdots<x_n$. 注意到
$$\mathbb{P}(X_{(1)}<x_1,\cdots,X_{(n)}<x_n)=\sum_{\pi}\mathbb{P}(X_{\pi_1}<x_1,\cdots,X_{\pi_n}<x_n,X_{\pi_1}<\cdots<X_{\pi_n}),$$
即可. 其中 $(\pi_1,\cdots\pi_n)$ 为 $(1,2,\cdots,n)$ 的排列.

(3) 注意到
$$\int\cdots\int_{a<x_1<\cdots<x_n<b}f(x_1)\cdots f(x_n)\mathrm{d}x_1\cdots\mathrm{d}x_n=\frac{1}{n!}(F(b)-F(a))^n.$$
我们有
$$\begin{aligned}g_i(x)=&\int\cdots\int_{-\infty<x_1<\cdots<x_n<+\infty}n!f(x_1)\cdots f(x_{i-1})f(x)f(x_{i+1})\cdots f(x_n)\mathrm{d}x_1\cdots\mathrm{d}x_{i-1}\mathrm{d}x_{i+1}\mathrm{d}x_n\\ =&n!\int\cdots\int_{-\infty<x_1<\cdots<x_{i-1}<x}f(x_1)\cdots f(x_{i-1})f(x)\mathrm{d}x_1\cdots\mathrm{d}x_{i-1}\\ &\times\int\cdots\int_{x<x_{i+1}<\cdots<x_n<+\infty}f(x_{i+1})\cdots f(x_n)\mathrm{d}x_{i+1}\cdots\mathrm{d}x_n\end{aligned}$$

$$=\frac{n!}{(i-1)!(n-i)!}(F(x))^{i-1}(1-F(x))^{n-i}f(x).$$

及

$$\begin{aligned}
g_{i,j}(x,y)=&n!\int\cdots\int_{-\infty<x_1<\cdots<x_{i-1}<x}f(x_1)\cdots f(x_{i-1})f(x)\mathrm{d}x_1\cdots\mathrm{d}x_{i-1}\\
&\times\int\cdots\int_{x<x_{i+1}<\cdots<x_{j-1}<y}f(x_{i+1})\cdots f(x_{j-1})f(y)\mathrm{d}x_{i+1}\cdots\mathrm{d}x_{j-1}\\
&\times\int\cdots\int_{y<x_{j+1}<\cdots<x_n<+\infty}f(x_{j+1})\cdots f(x_n)\mathrm{d}x_{j+1}\cdots\mathrm{d}x_n\\
=&\frac{n!}{(i-1)!(j-i-1)!(n-j)!}(F(x))^{i-1}(F(y)-F(x))^{j-i-1}(1-F(y))^{n-j}f(x)f(y).
\end{aligned}$$

(4) **提示**: 作变量代换: $Z=X_{(1)},R=X_{(n)}-X_{(1)}$ 即可, 这里 $|J|=1$, 代换后求 R 的边缘密度即可. 答案:

$$g_R(r)=\begin{cases}n(n-1)r^{n-2}(1-r), & 0<r<1,\\ 0, & \text{else.}\end{cases}$$

□

例 2.6 (指数分布)**.** $X\sim\mathrm{Exp}(\lambda),\lambda>0$, 密度函数

$$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x\geq 0,\\ 0, & x<0.\end{cases}$$

分布函数 $F(x)=1-e^{-\lambda x},x\geq 0$. 自行验证: $\mathbb{E}[X]=\dfrac{1}{\lambda},Var(X)=\dfrac{1}{\lambda^2}$.

(1) 若 X 是非负连续性随机变量, 则 X 是指数分布 $\Longleftrightarrow$ X 无记忆性.

(2) 设 $X_1,\cdots X_n$ 相互独立且分别服从参数为 $\lambda_1,\cdots\lambda_n$ 指数分布, 求 $Y=\min\{X_1,\cdots,X_n\}$ 的分布.

(3) 设 $X_1,\cdots X_n$ i.i.d $\sim\mathrm{Exp}(\lambda)$ 定义 $Y=X_1+\cdots+X_n$, 则 Y 的密度函数是

$$g(x)=\frac{\lambda^n}{\Gamma(n)}x^{n-1}\mathrm{e}^{-\lambda x},x\geqslant 0$$

其中 $\Gamma(n)=(n-1)!$. 我们记 Y 服从 Γ 分布 $\Gamma(n,\lambda)$. 请读者自行验证 Γ 分布具有可加性: 若 $Z_1\sim\Gamma(n,\lambda),Z_2\sim\Gamma(m,\lambda)$, Z_1,Z_2 独立, 则有 $Z_1+Z_2\sim\Gamma(n+m,\lambda)$ (直接计算或通过特征函数说明).

解. (1) $\Longrightarrow$: 若 $X\sim\mathrm{Exp}(\lambda)$, 则对 $s,t>0$ 有

$$\mathbb{P}(X>t+s\mid X>t)=\frac{\mathbb{P}(X>t+s)}{\mathbb{P}(X>t)}=\frac{\mathrm{e}^{-\lambda(t+s)}}{\mathrm{e}^{-\lambda t}}=e^{-\lambda s}=\mathbb{P}(X>s);$$

$\Longleftarrow$: 记 $G(x)=\mathbb{P}(X>x)$, 则有

$$G(s+t)=G(s)G(t).$$

其中 $s,t\geqslant 0$. 对 $\forall n\in\mathbb{N}^*$ 及 $x\in\mathbb{R}^*$, 有 $G(nx)=(G(x))^n$. 取 $x=\dfrac{1}{n}$, 并记 $a=G(1)\geqslant 0$, 则有

$$G\left(\frac{1}{n}\right)=a^{\frac{1}{n}}.$$

因此, 对任意正整数 m,n, 有

$$G\left(\frac{m}{n}\right)=\left(G\left(\frac{1}{n}\right)\right)^m=a^{\frac{m}{n}}.$$

因此 $G(x)=a^x$ 对所有正有理数成立, 利用 $G(x)$ 的连续性 (或单调性), 则 $G(x)=a^x$ 对 $\forall x\geqslant 0$ 均成立. 因为 $G(x)\in[0,1]$, 且 $\lim\limits_{x\to+\infty}G(x)=0$, 故 $a\in(0,1)$, 可写为 $a=\mathrm{e}^{-\lambda}$, 其中 $\lambda>0$. 因此

$$F(x)=1-G(x)=1-\mathrm{e}^{-\lambda x}.$$

(2) X_i 的分布函数 $F_i(x)=1-\mathrm{e}^{-\lambda_i x}$, 类似**例 2.5(1)**, 可得 Y 的分布函数

$$F_Y(x)=1-\prod_{i=1}^n(1-F_i(x))=1-\mathrm{e}^{-(\lambda_1+\cdots+\lambda_n)x},\ x>0.$$

因此 $Y\sim\mathrm{Exp}(\lambda_1+\cdots+\lambda_n)$.

(3) 可通过归纳计算得到 (自行验证), 或作变量代换 $Y_k=X_1+\cdots+X_k$, $k=1,\cdots,n$, 则有

$$\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix}=\begin{pmatrix}1&&&\\1&1&&\\\vdots&\vdots&\ddots&\\1&1&\cdots&1\end{pmatrix}\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}:=A\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}.$$

$$|J|=\det\left(\frac{\partial x_i}{\partial y_j}\right)_{ij}=\det(A^{-1})=1.$$

而 $(X_1,\cdots X_n)$ 的联合密度是

$$f(x_1,\cdots x_n)=\lambda^n\mathrm{e}^{-\lambda(x_1+\cdots+x_n)},\ x_1,\cdots,x_n>0.$$

故 $(Y_1,\cdots Y_n)$ 的联合密度为

$$g(y_1,\cdots y_n)=\lambda^n\mathrm{e}^{-\lambda y_n},\ y_n\geqslant y_{n-1}\geqslant\cdots\geqslant y_1\geqslant 0.$$

所以 Y_n 的联合密度

$$\begin{aligned}g_n(y_n)&=\int_0^{y_n}\int_0^{y_{n-1}}\cdots\int_0^{y_2}g(y_1,\cdots y_n)\mathrm{d}y_1\cdots\mathrm{d}y_{n-1}=\lambda^n\mathrm{e}^{-\lambda y_n}\int_0^{y_{n-1}}\cdots\int_0^{y_2}1\ \mathrm{d}y_1\cdots\mathrm{d}y_{n-1}\\&=\frac{\lambda^n}{(n-1)!}{y_n}^{n-1}\mathrm{e}^{-\lambda y_n}.\end{aligned}$$

□

例 2.7 (正态分布 (高斯分布)). $X\sim N(\mu,\sigma^2)$, 分布函数记为 $F(x)$, 密度函数

$$f(x)=\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{1}{2\sigma^2}(x-\mu)^2},\ -\infty<x<+\infty,\ \mu\in\mathbb{R},\ \sigma>0.$$

自行验证: $\mathbb{E}[X]=\mu, Var(X)=\sigma^2$.

(1) 请读者自行验证正态分布具有**可加性**: 若 $Z_1 \sim N(\mu_1, \sigma_1^2), Z_2 \sim N(\mu_2, \sigma_2^2)$, Z_1, Z_2 独立, 则有 $Z_1 + Z_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$ (直接计算或通过特征函数说明).

(2) $\mu = 0, \sigma = 1$. 记 $\gamma_n = \mathbb{E}[X^n]$, 则有

$$\gamma_n = \begin{cases} 0, & n = 2k-1, \\ (2k-1)!!, & n = 2k. \end{cases}$$

由此可知, 正态分布的任意阶矩均有限, 且奇阶矩为零.

(3) **Mill's ratio** $\mu = 0, \sigma = 1$, 利用 $f'(x) + xf(x) = 0$, 证明:

$$\frac{1}{x} - \frac{1}{x^3} < \frac{1 - F(x)}{f(x)} < \frac{1}{x}.$$

更一般地, 请读者自证:

$$\frac{1}{x} - \frac{1}{x^3} + \frac{3}{x^5} - \frac{15}{x^7} < \frac{1 - F(x)}{f(x)} < \frac{1}{x} - \frac{1}{x^3} + \frac{3}{x^5}.$$

(4) $\mu = 0, \sigma = 1, a > 0$, 证明: 当 $x \to +\infty$ 时, 有

$$\mathbb{P}\left(X > x + \frac{a}{x} \Big| X > x\right) \to \mathrm{e}^{-a}.$$

解. (2) 奇阶矩为零显然, 因为函数 $x^{2k-1}f(x)$ 是奇函数. 对于偶阶矩, 令 $t = \dfrac{x^2}{2}$, 则有

$$\mathbb{E}[X^{2k}] = \int_{-\infty}^{\infty} \frac{x^{2k}}{\sqrt{2\pi}} \mathrm{e}^{-\frac{x^2}{2}} \mathrm{d}x = \int_0^{\infty} \frac{2^{k-1} t^{k-\frac{1}{2}}}{\sqrt{\pi}} \mathrm{e}^{-t} \mathrm{d}t = \frac{2^{k-1}}{\sqrt{\pi}} \Gamma(k + \frac{1}{2}) = (2k-1)!!.$$

(3) 反复利用分部积分, 并利用 $f'(x) + xf(x) = 0$, 有

$$1 - F(x) = \int_x^{+\infty} f(u) \mathrm{d}u = -\int_x^{+\infty} \frac{f'(u)}{u} \mathrm{d}u = \frac{f(x)}{x} + \int_x^{+\infty} \frac{f'(u)}{u^3} \mathrm{d}u = \frac{f(x)}{x} - \frac{f(x)}{x^3} - \int_x^{+\infty} \frac{3f'(u)}{u^5} \mathrm{d}u.$$

(4) 利用 (3) 可知,

$$\mathbb{P}\left(X > x + \frac{a}{x} \Big| X > x\right) = \frac{1 - F\left(x + \frac{a}{x}\right)}{1 - F(x)} = (1 + o(1)) \frac{\mathrm{e}^{-\frac{1}{2}(x + \frac{a}{x})^2}}{\mathrm{e}^{-\frac{1}{2}x^2}} \xrightarrow{x \to +\infty} \mathrm{e}^{-a}.$$

□

2.3 正态分布下的样本均值与样本方差

两个统计量: 从总体中抽取样本 $X_1, \cdots, X_n$

样本均值: $\overline{X} = \dfrac{1}{n} \displaystyle\sum_{k=1}^n X_k$

样本方差: $S^2 = \dfrac{1}{n-1} \displaystyle\sum_{k=1}^n (X_k - \overline{X})^2$

例 2.8. $X_1, \cdots, X_n$ i.i.d X, $\mathbb{E}[X] = \mu, Var(X) = \sigma^2$, 则 $\mathbb{E}[\overline{X}] = \mu, Var(\overline{X}) = \dfrac{\sigma^2}{n}, \mathbb{E}[S^2] = \sigma^2$(用来估计 μ 和 σ^2)

证明. 只证明 $\mathbb{E}[S^2]=\sigma^2$.

$$\begin{aligned}\mathbb{E}[S^2]&=\frac{1}{n-1}\sum_{i=1}^{n}\mathbb{E}[(x_i-\mu-(\overline{X}-\mu))^2]=\frac{1}{n-1}\sum_{i=1}^{n}(Var(X_i)+Var(\overline{X})-2\mathbb{E}[(X_i-\mu)(\overline{X}-\mu)])\\&=\frac{1}{n-1}\left(n\sigma^2+n\cdot\frac{1}{n}\sigma^2-2n\cdot\frac{1}{n}\sigma^2\right)\\&=\sigma^2.\end{aligned}$$

□

定义 2.1. 当 X 有密度函数

$$f(x)=\frac{1}{2^{\frac{d}{2}}\Gamma\left(\frac{d}{2}\right)}x^{\frac{d}{2}-1}e^{-\frac{1}{2}x},\ x>0.$$

则 X 服从 d 个自由度的 卡方分布, 记 $X\sim\chi^2(d)$.

注. $\chi^2(d)=\Gamma\left(\frac{d}{2},\frac{1}{2}\right)$.

引理 2.2. $Y_1,\cdots,Y_d$ 独立同 $N(0,1)$, $X=\sum_{i=1}^{d}Y_i^2$, 则 $X\sim\chi^2(d)$.

证明. $(Y_1,\cdots,Y_d)$ 联合密度 $f(y_1,\cdots,y_d)=\frac{1}{(\sqrt{2\pi})^d}e^{-\frac{1}{2}\sum_{j=1}^{d}y_j^2}$, 因此

$$F_X(x)=\mathbb{P}\left(\sum_{j=1}^{d}Y_i^2\le x\right)=\int_{\sum y_i^2\le x}f(y_1,\cdots,y_d)\mathrm{d}y_1\cdots\mathrm{d}y_d.$$

极坐标换元后有:

$$\begin{aligned}F_X(x)&=\mathrm{C}_d\int_0^{\sqrt{x}}r^{d-1}e^{-\frac{1}{2}r^2}\mathrm{d}r\\&=\frac{\mathrm{C}_d}{2}\int_0^{x}r^{\frac{d}{2}-1}e^{-\frac{1}{2}r}\mathrm{d}r.\end{aligned}$$

利用 $F_X(\infty)=1$, 得: $\frac{\mathrm{C}_d}{2}=\frac{1}{2^{\frac{d}{2}}\Gamma\left(\frac{d}{2}\right)}$. □

注. 卡方分布具有可加性, 其证明已留为作业.

定理 2.3. $X_1,\cdots,X_n$ 独立同 $N(\mu,\sigma^2)$, 则

(1) $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i\sim N(\mu,\frac{\sigma^2}{n})$.

(2) $\frac{n-1}{\sigma^2}S^2\sim\chi^2(n-1)$.

(3) $\overline{X}$ 与 S^2 独立.

证明. 令 $Y_i=\frac{1}{\sigma}(X_i-\mu)$, 则 $Y_i\sim N(0,1)$, 且 $Y_1,\cdots,Y_n$ 相互独立, 显见

$$\overline{Y}=\frac{1}{n}\sum_{i=1}^{n}Y_i=\frac{1}{\sigma}(\overline{X}-\mu).$$

$$\sum_{i=1}^n (Y_i-\overline{Y})^2=\frac{1}{\sigma^2}\sum_{i=1}^n (X_i-\overline{X})^2=\frac{n-1}{\sigma^2}S^2.$$

$(Y_1,\cdots,Y_n)$ 密度 $f(y)=\dfrac{1}{(\sqrt{2\pi})^n}e^{-\frac{1}{2}\sum_{i=1}^n y_i^2}$, 取正交阵

$$A=\begin{pmatrix} \frac{1}{\sqrt{n}} & * & * & * \\ \vdots & * & * & * \\ \vdots & * & * & * \\ \frac{1}{\sqrt{n}} & * & * & * \end{pmatrix},$$

令 $(Z_1,\cdots,Z_n)=(Y_1,\cdots,Y_n)A$, 则 $Z_1=\sqrt{n}\overline{Y}$, 且 $Z_1,\cdots,Z_n$ 独立同 $N(0,1)$, 因为 $\vec{Y}\sim N(0,I_n)$ 知 $\vec{Z}\sim N(0,A^TI_nA)=N(0,I_n)$. 又可知

$$Z_1^2+\cdots+Z_n^2=Y_1^2+\cdots+Y_n^2,$$

有

$$\begin{aligned}\sum_{i=2}^n Z_i^2&=\sum_{i=1}^n Y_i^2-n\overline{Y}^2=\sum_{i=1}^n (Y_i-\overline{Y}+\overline{Y})^2-n\overline{Y}^2\\&=\sum_{i=1}^n (Y_i-\overline{Y})^2=\frac{n-1}{\sigma^2}S^2.\end{aligned}$$

因此 $\overline{Y}\sim N(0,\frac{1}{n})$, $\dfrac{n-1}{\sigma^2}S^2\sim\chi^2(n-1)$, 且 $\overline{X}$ 与 S^2 独立. □

3 补充习题

1 Grimmett 4.3.3, 4.6.7, 4.6.8, 4.7.2, 4.7.7(4.14.18), 4.14.14, 4.14.19, 4.14.27, 4.14.28, 4.14.29, 4.14.33, 4.14.39, 4.14.46 (其中部分题与精选题重合)

2 若 X,Y 独立且服从 $N(0,1)$, 求 $U=X^2+Y^2$ 与 $V=\dfrac{X}{Y}$ 的密度函数, 并证明它们是独立的.

提示. 可以计算 $J^{-1}=-2(v^2+1)$. 求出联合密度后可直接得到独立性, 答案:

$$f_U(u)=\frac{1}{2}\mathrm{e}^{-\frac{u}{2}},\ u>0,\quad f_V(v)=\frac{1}{\pi(1+v^2)},v\in\mathbb{R}.$$

3 若气体分子的速度是随机向量 $V=(X,Y,Z)$, 各分量相互独立且服从 $N(0,\sigma^2)$. 证明: $S=\sqrt{X^2+Y^2+Z^2}$ 服从 **Maxwell 分布律**:

$$p(s)=\sqrt{\frac{2}{\pi}}\frac{s^2}{\sigma^3}\exp\left(-\frac{s^2}{2\sigma^2}\right),s>0.$$

提示. 利用球坐标换元计算即可.

4 对连续型随机变量 X, 定义实数 m 为 X 的分布函数 F 的**中位数**, 若

$$F(m-0) \leqslant \frac{1}{2} \leqslant F(m) \Longleftrightarrow \mathbb{P}(X \leqslant m) \geqslant \frac{1}{2} \text{ 且 } \mathbb{P}(X \geqslant m) \geqslant \frac{1}{2}.$$

(1) **(本小问第 2 次习题课讲义已证)** 证明分布函数至少有一个中位数, 且中位数集合构成一个闭区间.

(2) 若 $\mathbb{E}[|X|] < \infty$, 证明:

$$\mathbb{E}[|X-m|] = \inf_{x\in\mathbb{R}} \mathbb{E}[|X-x|]$$

当且仅当 m 是 X 的中位数.

(3) 若 $\mathbb{E}[X^2] < \infty$, m 是 X 的中位数. 证明:

$$|\mathbb{E}[X] - m| \leqslant \sqrt{Var(X)}.$$

解. (2) **提示**: 用 Fubini 定理证明: 对 $\forall a, b \in \mathbb{R}$, 有

$$\mathbb{E}[|X-b|] - \mathbb{E}[|X-a|] = \int_a^b (\mathbb{P}(X \leqslant x) - \mathbb{P}(X \geqslant x))\mathrm{d}x$$

(3) 利用 Cauchy-Schwarz 不等式, 结合 (2) 可得

$$|\mathbb{E}[X] - m| = |\mathbb{E}[X-m]| \leqslant \mathbb{E}[|X-m|] \leqslant \mathbb{E}[|X-\mathbb{E}[X]|] \leqslant \sqrt{Var(X)}.$$

□