Chapter 12: $\mathbb{R}^n$ 上的概率分布

Latest Update: 2025 年 1 月 1 日

Exercise #12. 1. 证明:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{\frac{-\left(x^2+y^2\right)}{2\sigma^2}}dxdy = 2\pi\sigma^2$$

因此, $\frac{1}{2\pi\sigma^2}e^{-\left(x^2+y^2\right)/2\sigma^2}$ 是一个概率密度函数.

证明. 做三角变换. 令

$$\begin{cases} x = r\cos(\theta), \\ y = r\sin(\theta). \end{cases}$$

其中, $r \geq 0$, $\theta \in [0, 2\pi)$. 于是逆变换为

$$\begin{cases} r = \sqrt{x^2+y^2}, \\ \theta = \arctan\left(\frac{y}{x}\right). \end{cases}$$

计算雅可比行列式:

$$\begin{aligned} J = \frac{\partial(x,y)}{\partial(r,\theta)} &= \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial y}{\partial r} \\ \frac{\partial x}{\partial \theta} & \frac{\partial y}{\partial \theta} \end{vmatrix} \\ &= \begin{vmatrix} \cos(\theta) & \sin(\theta) \\ -r\sin(\theta) & r\cos(\theta) \end{vmatrix} \\ &= r. \end{aligned}$$

于是函数的积分是

$$\begin{aligned} &\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{\frac{-\left(x^2+y^2\right)}{2\sigma^2}}dxdy \\ =&\int_0^{2\pi} d\theta \int_0^{\infty} \exp\left(-\frac{r^2}{2\sigma^2}\right) r\,dr \\ =&2\pi\int_0^{\infty} \exp\left(-\frac{r^2}{2\sigma^2}\right) r\,dr \\ =&2\pi\sigma^2.aqaaq \end{aligned}$$

从而 $\frac{1}{2\pi\sigma^2}e^{-\left(x^2+y^2\right)/2\sigma^2}$ 是一个概率密度函数. □

Exercise #12. 2. 假设一个联合密度可以分解为: $f_{(X,Y)}(x,y) = g(x)h(y)$. 求出 $f_X(x)$ 和 $f_Y(y)$.

证明. 根据定义,

$$f_X(x) = \int_{-\infty}^{\infty} f_{(X,Y)}(x,y)dy = \int_{-\infty}^{\infty} g(x)h(y)dy = g(x)\int_{-\infty}^{\infty} h(y)\,dy.$$

同理, $f_Y(y) = h(y)\int_{-\infty}^{\infty} g(x)\,dx$. 由于

$$\iint_{\mathbb{R}^2} f_{(X,Y)}(x,y)dxdy = 1,$$

从而 $\int_{-\infty}^{\infty} g(x)\,dx \int_{-\infty}^{\infty} h(y)\,dy = 1$, 这表明, 如果 $f_{(X,Y)}(x,y) = g(x)h(y)$, 可以写成因子分解的形式, 那么 X, Y 分别的密度函数是

$$f_X(x) = \frac{g(x)}{\int_{-\infty}^{\infty} g(x)\,dx},$$
$$f_Y(y) = \frac{h(y)}{\int_{-\infty}^{\infty} h(y)\,dy}.$$

还有 $X \perp\!\!\!\perp Y$. □

Exercise #12. 3. 设 (X,Y) 有联合密度

$$\begin{aligned} f(x,y) = &\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} \\ &\times \exp\left(-\frac{1}{2\left(1-r^2\right)}\left\{\frac{\left(x-\mu_1\right)^2}{\sigma_1^2} - \frac{2r\left(x-\mu_1\right)\left(y-\mu_2\right)}{\sigma_1\sigma_2} + \frac{\left(y-\mu_2\right)^2}{\sigma_2^2}\right\}\right) \end{aligned}$$

求出 $f_{X=x}(y)$.

证明. 按目前条件密度的定义,

$$f_{X=x}(y) = \frac{f(x,y)}{f_X(x)}.$$

而

$$\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} = \frac{1}{\sqrt{2\pi\sigma_1^2}}\frac{1}{\sqrt{2\pi\sigma_2^2(1-r^2)}}$$

以及

$$
\begin{aligned}
&\exp\left\{-\frac{1}{2(1-r^2)}\left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2-2r\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right)+\left(\frac{y-\mu_2}{\sigma_2}\right)^2\right]\right\}\\
=&\exp\left\{-\frac{1}{2(1-r^2)}\left[(1-r^2)\left(\frac{x-\mu_1}{\sigma_1}\right)^2+r^2\left(\frac{x-\mu_1}{\sigma_1}\right)^2-2r\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right)+\left(\frac{y-\mu_2}{\sigma_2}\right)^2\right]\right\}\\
=&\exp\left\{-\frac{1}{2\sigma_1^2}(x-\mu_1)^2\right\}\exp\left\{-\frac{1}{2(1-r^2)}\left[\left(\frac{y-\mu_2}{\sigma_2}\right)-r\left(\frac{x-\mu_1}{\sigma_1}\right)\right]^2\right\}\\
=&\underbrace{\exp\left\{-\frac{1}{2\sigma_1^2}(x-\mu_1)^2\right\}}_{\frac{1}{\sqrt{2\pi\sigma_1^2}}}\cdot\underbrace{\exp\left\{-\frac{1}{2(1-r^2)\sigma_2^2}\left[(y-\mu_2)-r\frac{\sigma_2}{\sigma_1}(x-\mu_1)\right]^2\right\}}_{\frac{1}{\sqrt{2\pi\sigma_2^2(1-r^2)}}}
\end{aligned}
$$

接下来, 我们根据上述分解式计算二元正态的边际密度函数.

$$
\begin{aligned}
f_X(x)=&\int_{-\infty}^{\infty}f(x,y)\,dy\\
=&(2\pi\sigma_1^2)^{\frac{1}{2}}\exp\left\{-\frac{1}{2\sigma_1^2}(x-\mu_1)^2\right\}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi\sigma_2^2(1-r^2)}}\exp\left\{-\frac{1}{2(1-r^2)\sigma_2^2}\left[y-\left\{\mu_2+r\frac{\sigma_2}{\sigma_1}(x-\mu_1)\right\}\right]^2\right\}dy\\
=&(2\pi\sigma_1^2)^{\frac{1}{2}}\exp\left\{-\frac{1}{2\sigma_1^2}(x-\mu_1)^2\right\}
\end{aligned}
$$

于是

$$
f_{X=x}(y)=\frac{f(x,y)}{f_X(x)}=\frac{1}{\sqrt{2\pi\sigma_2^2(1-r^2)}}\exp\left\{-\frac{1}{2(1-r^2)\sigma_2^2}\left[y-\left\{\mu_2+r\frac{\sigma_2}{\sigma_1}(x-\mu_1)\right\}\right]^2\right\}.
$$

服从均值为 $\mu_2+r\dfrac{\sigma_2}{\sigma_1}(x-\mu_1)$, 方差为 $\sigma_2^2(1-r^2)$ 的正态分布. □

注. 上述推导也可以用多元正态分布的推导进行.

$$
\begin{pmatrix}X\\Y\end{pmatrix}\sim\mathcal{N}_2\left(\begin{pmatrix}\mu_1\\\mu_2\end{pmatrix},\begin{bmatrix}\sigma_1^2&r\sigma_1\sigma_2\\r\sigma_1\sigma_2&\sigma_2^2\end{bmatrix}\right)
$$

根据多元正态分布的线性组合还是正态分布, 取 $B=(1\quad 0)$, 有 $X=B\begin{pmatrix}X\\Y\end{pmatrix}\sim\mathcal{N}_1(\mu_1,\sigma_1^2)$.

$$
Y|X\sim\mathcal{N}_1\left(\mu_{2\cdot1},\Sigma_{22\cdot1}\right).
$$

其中

$$
\mu_{2\cdot1}=\mu_2+\Sigma_{21}\Sigma_{11}^{-1}(x-\mu_1),\quad \Sigma_{22\cdot1}=\Sigma_{22}-\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}.
$$

Exercise #12. 4. 令 $\rho_{X,Y}$ 表示 (X,Y) 的相关系数. 设 $a>0,c>0,b\in\mathbb{R}$. 证明:

$$
\rho_{aX+b,cY+b}=\rho_{X,Y}
$$

这表明, 相关系数不受测量尺度 (线性变换) 的影响.

证明.

$$\begin{aligned}\rho_{aX+b,cY+b}=&\frac{\operatorname{Cov}(aX+b,cY+d)}{\sqrt{\operatorname{Var}(aX+b)\operatorname{Var}(cY+d)}}\\=&\frac{ac\operatorname{Cov}(X,Y)}{\sqrt{a^2c^2\operatorname{Var}(X)\operatorname{Var}(Y)}}\\=&\frac{\operatorname{Cov}(X,Y)}{\sqrt{\operatorname{Var}(X)\operatorname{Var}(Y)}}\\=&\rho_{X,Y}.\end{aligned}$$

□

Exercise #12. 5. 若 $a\neq 0$, 证明:

$$\rho_{X,aX+b}=\frac{a}{|a|},$$

于是, 若 $Y=aX+b$ 是 X 的非常数仿射变换, 则 $\rho_{X,Y}=\pm 1$.

证明.

$$\begin{aligned}\rho_{X,aX+b}=&\frac{\operatorname{Cov}(X,aX+b)}{\sqrt{\operatorname{Var}(X)\operatorname{Var}(aX+b)}}\\=&\frac{a\operatorname{Var}(X)}{\sqrt{\operatorname{Var}(X)\operatorname{Var}(aX+b)}}\\=&\frac{a}{|a|}.\end{aligned}$$

□

Exercise #12. 6. 设 X,Y 方差有限, 令

$$Z=\left(\frac{1}{\sigma_Y}\right)Y-\left(\frac{\rho_{X,Y}}{\sigma_X}\right)X.$$

证明: $\sigma_Z^2=1-\rho_{X,Y}^2$, 并推导若 $\rho_{X,Y}=\pm 1$, 则 Y 是 X 的非常数仿射变换.

证明.

$$\begin{aligned}\sigma_Z^2=&\operatorname{Var}\left(\left(\frac{1}{\sigma_Y}\right)Y-\left(\frac{\rho_{X,Y}}{\sigma_X}\right)X\right)\\=&\left(\frac{1}{\sigma_Y}\right)^2\operatorname{Var}(Y)+\left(\frac{\rho_{X,Y}}{\sigma_X}\right)^2\operatorname{Var}(X)-2\left(\frac{1}{\sigma_Y}\right)\left(\frac{\rho_{X,Y}}{\sigma_X}\right)\operatorname{Cov}(X,Y)\\=&1+\rho_{X,Y}^2-2\rho_{X,Y}\\=&1-\rho_{X,Y}^2.\end{aligned}$$

若 $\rho_{X,Y}=\pm 1$, 则 $\sigma_Z^2=0$, 从而 Z 是几处常数 c,

$$\left(\frac{1}{\sigma_Y}\right)Y-\left(\frac{\rho_{X,Y}}{\sigma_X}\right)X=c$$

从而

$$Y=\sigma_Yc+\rho_{X,Y}\frac{\sigma_Y}{\sigma_X}X.$$

$\square$

Exercise #12. 7. 令 (X,Y) 是单位圆上的均匀分布. 即

$$f_{(X,Y)}(x,y)=\begin{cases}\dfrac{1}{\pi}\ if\ x^2+y^2\leq 1\\ 0\ if\ x^2+y^2>1\end{cases}$$

求出 $R=\sqrt{X^2+Y^2}$ 的分布.

证明. 设 F_R 为 R 的分布函数. 当 $r\leq 0$ 时, $F_R(r)=0$; 当 $r\geq 1$ 时, $F_R(r)=1$. $\forall r\in(0,1)$,

$$\begin{aligned}F_R(r)=&P\{R\leq r\}\\ =&r^2.\end{aligned}$$

于是密度函数为

$$f_R(r)=2r\mathbb{I}_{(0,1)}(r)$$

$\square$

Exercise #12. 8. 设 (X,Y) 有密度函数 $f(x,y)$. 求出 $Z=X+Y$ 的密度函数.

证明. 使用卷积公式,

$$F_Z(z)=P\{Z\leq z\}=\iint_{x+y\leq z}f(x,y)\,dx\,dy=\int_{-\infty}^{\infty}\int_{-\infty}^{z-x}f(x,y)\,dy\,dx.$$

令 $y=t-x$, 再运用 Fubini 定理,

$$\begin{aligned}F_Z(z)&=\int_{-\infty}^{\infty}\int_{-\infty}^{z}f(x,t-x)\,dt\,dx\\ &=\int_{-\infty}^{z}\int_{-\infty}^{\infty}f(x,t-x)\,dx\,dt\end{aligned}$$

根据密度函数的定义,

$$f_Z(z)=\int_{-\infty}^{\infty}f(x,z-x)\,dx=\int_{-\infty}^{\infty}f(z-y,y)\,dy.$$

$\square$

Exercise #12. 9. 设 X 服从 $\mu=0$, $\sigma^2<\infty$ 的正态分布. 设 Θ 服从 $(0,\pi)$ 上的均匀分布, 即密度函数 $f(\theta)=\dfrac{1}{\pi}\mathbb{I}_{(0,\pi)}(\theta)$. 假设 X 和 Θ 是独立的. 求出 $Z=X+a\cos(\Theta)$ 的密度函数.

注. 这在电子工程中很有用.

证明. Θ 的密度函数为

$$f(\theta) = \frac{1}{\pi}\mathbb{I}_{(0,\pi)}(\theta).$$

令 $Y = a\cos(\Theta)$, 取 $g(\theta) = a\cos(\theta)$, 在 $(0,\pi)$ 上是单射, 且 $g^{-1}(y) = \arccos\left(\frac{y}{a}\right)$. 导数为

$$\frac{d}{dy}\arccos\left(\frac{y}{a}\right) = -\frac{1}{\sqrt{a^2-y^2}}.$$

于是 Y 的密度函数是

$$f_Y(y) = f(\arccos\left(\frac{y}{a}\right))\left|\frac{d}{dy}\arccos\left(\frac{y}{a}\right)\right|_+ = \frac{1}{\pi\sqrt{a^2-y^2}}\mathbb{I}_{(-a,a)}(y).$$

此外, X 的密度函数是

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{x^2}{2\sigma^2}\right\}.$$

根据独立性,

$$f_{X,Y}(x,y) = f_X(x)f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{x^2}{2\sigma^2}\right\}\frac{1}{\pi\sqrt{a^2-y^2}}\mathbb{I}_{(-a,a)}(y).$$

根据上一题的卷积公式,

$$\begin{aligned}
f_Z(z) &= \int_{-\infty}^{\infty} f_X(z-y)f_Y(y)\,dy \\
&= \int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{(z-y)^2}{2\sigma^2}\right\}\frac{1}{\pi\sqrt{a^2-y^2}}\mathbb{I}_{(-a,a)}(y)\,dy \\
&= \frac{1}{\sqrt{2\pi}\sigma}\int_0^{\pi}\exp\left\{-\frac{(z-a\cos(w))^2}{2\sigma^2}\right\}\frac{1}{\pi\sqrt{a^2-a^2\cos^2(w)}}a\sin(w)dw \\
&= \frac{1}{\pi\sigma\sqrt{2\pi}}\int_0^{\pi}\exp\left\{-\frac{(z-a\cos(w))^2}{2\sigma^2}\right\}dw.
\end{aligned}$$

□

Exercise #12. 10. 设 X, Y 是独立的, 假设 $Z = g(X), W = h(Y)$, 其中 g 和 h 是单射, 可微函数. 求出 (Z,W) 的联合密度 $f_{Z,W}(z,w)$.

证明. 由于 X, Y 独立, g, h 可微, 从而可测, 于是 Z, W 也是独立的.

$$\begin{aligned}
f_{Z,W}(z,w) &= f_Z(z)f_W(w) \\
&= f_X(g^{-1}(z))f_Y(h^{-1}(w))\left|\frac{\partial g^{-1}}{\partial z}\right|_+\left|\frac{\partial h^{-1}}{\partial w}\right|_+.
\end{aligned}$$

□

Exercise #12. 11. 设 X, Y 是独立的, 且都服从 $\mu = 0$, $\sigma^2 < \infty$ 的正态分布. 令

$$Z = \sqrt{X^2+Y^2}\ ,\ W = \arctan\left(\frac{X}{Y}\right),\quad -\frac{\pi}{2} < W \le \frac{\pi}{2}.$$

证明 Z 服从 *Rayleigh* 分布, W 服从 $\left(-\frac{\pi}{2},\frac{\pi}{2}\right)$ 上的均匀分布. Z 和 W 独立.

证明. 由于 (X,Y) 的联合密度函数为

$$f_{(X,Y)}(x,y) = \frac{1}{2\pi\sigma^2}\exp\left\{-\frac{x^2+y^2}{2\sigma^2}\right\},$$

由于

$$\begin{cases} Z = \sqrt{X^2+Y^2} \in [0,\infty), \\ W = \arctan\left(\dfrac{X}{Y}\right) \in \left(-\dfrac{\pi}{2}, \dfrac{\pi}{2}\right), \end{cases}$$

于是可以分两种情况考虑. 当 $Y>0$ 时, $Y = Z\cos(W)$; 当 $Y<0$ 时, $Y = -Z\cos(W)$. 定义

$$\begin{aligned} S_0 &= \{(x,y) : y = 0\}, \\ S_1 &= \{(x,y) : y > 0\}, \\ S_2 &= \{(x,y) : y < 0\}. \end{aligned}$$

于是 S_0, S_1, S_2 构成了对平面 $\mathbb{R}^2$ 的划分. 而且 $m_2(S_0) = 0$. 对于 S_1, S_2,

$$g(x,y) = \left(\sqrt{x^2+y^2}, \arctan\left(\frac{x}{y}\right)\right)$$

是单射.(因为是二维的, 只要 x,y 中有一个正负号识别出就可以.) 于是在 S_1 上, g 可逆, 逆变换为

$$\begin{cases} X = Z\sin(W), \\ Y = Z\cos(W). \end{cases}$$

Jacobian 行列式为

$$\begin{vmatrix} \dfrac{\partial x}{\partial z} & \dfrac{\partial x}{\partial w} \\ \dfrac{\partial y}{\partial z} & \dfrac{\partial y}{\partial w} \end{vmatrix} = \begin{vmatrix} \sin(w) & z\cos(w) \\ \cos(w) & -z\sin(w) \end{vmatrix} = -z.$$

在 S_2 上, g 可逆, 逆变换为

$$\begin{cases} X = Z\sin(W), \\ Y = -Z\cos(W). \end{cases}$$

Jacobian 行列式为

$$\begin{vmatrix} \dfrac{\partial x}{\partial z} & \dfrac{\partial x}{\partial w} \\ \dfrac{\partial y}{\partial z} & \dfrac{\partial y}{\partial w} \end{vmatrix} = \begin{vmatrix} \sin(w) & z\cos(w) \\ -\cos(w) & z\sin(w) \end{vmatrix} = z.$$

从而, (Z,W) 的联合密度函数为

$$\begin{aligned} f_{Z,W}(z,w) &= \left(\frac{1}{2\pi\sigma^2}e^{\frac{-z^2}{2\sigma}}z + \frac{1}{2\pi\sigma^2}e^{\frac{-z^2}{2\sigma}}z\right)\mathbb{I}_{(-\frac{\pi}{2},\frac{\pi}{2})}(w)\mathbb{I}_{(0,\infty)}(z) \\ &= \frac{z}{\pi\sigma^2}\exp\left\{-\frac{z^2}{2\sigma^2}\right\}\mathbb{I}_{(0,\infty)}(z)\mathbb{I}_{(-\frac{\pi}{2},\frac{\pi}{2})}(w). \end{aligned}$$

于是根据习题 12.2 可得 Z 服从 Rayleigh 分布, W 服从 $\left(-\dfrac{\pi}{2}, \dfrac{\pi}{2}\right)$ 上的均匀分布. 且 Z 和 W 独立. □

注. 这一题要仔细分析 X,Y 的取值范围, 从而找出对应的单射区域.

Exercise #12. 12. 设 $(X_1, ..., X_n)$ 是独立随机变量. 定义

$$\begin{aligned} &Y_1 = \min(X_i; 1 \le i \le n) \\ &Y_2 = \textit{ second smallest of } X_1, \ldots, X_n \\ &\vdots \\ &Y_n = \textit{ largest of } X_1, \ldots, X_n. \end{aligned}$$

则 $Y_1, ..., Y_n$ 也是随机变量, 且有 $Y_1 \le Y_2 \le ... \le Y_n$. 因此, $Y_1, ..., Y_n$ 是 $X_1, ..., X_n$ 的顺序统计量. 通常记 $Y_k = X_{(k)}$. 假设 X_i 独立同分布于相同的密度函数 f. 证明顺序统计量的联合密度函数为

$$f_{\left(X_{(1)},\ldots,X_{(n)}\right)}(y_1, \ldots, y_n) = \begin{cases} n! \prod_{i=1}^{n} f(y_i) & \textit{for } y_1 < y_2 < \ldots < y_n \\ 0 & \textit{otherwise} \end{cases}$$

证明. 先证明: 顺序统计量是随机变量. $\forall x \in \mathbb{R}$, 定义集合

$$\begin{aligned} A_m(x) =& \{\text{刚好有}m\text{个}X_i \le x\} \\ =& \bigcup_{\tau\in\sigma\{1,\ldots,n\}} \{X_{\tau_1} \le x, ..., X_{\tau_m} \le x, X_{\tau_{m+1}} > x, ..., X_{\tau_n} > x\}. \end{aligned}$$

其中 $\tau = (\tau_1, ..., \tau_n)$ 表示 $\{1, ..., n\}$ 的一个置换. 从而 $A_m(x) \in \mathcal{B}$. 于是,

$$\{Y_k \le x\} = \{X_{(k)} \le x\} = \bigcup_{m \ge k} A_m(x) \in \mathcal{B}.$$

因此, 顺序统计量是随机变量.

接下来说明联合分布.

一种方法是按定义. 根据全排列, $X_1, ..., X_n$ 的支集可以分割成 $n!$ 个区域. 这些集合映射到 $Y_1, ..., Y_n$ 的支集上, 即

$$\{(y_1, ..., y_n) : a < y_1 < y_2 < ... < y_n < b\}.$$

对于没种可能的排列映射到 Y 上, 根据行列式的定义, 每一个变换的 Jacobian 只能是 ± 1. 因此,

$$\begin{aligned} g(y_1, y_2, \cdots, y_n) &= \sum_{i=1}^{n!} |J_i| f(y_1) f(y_2) \cdots f(y_n) \\ &= \begin{cases} n! f(y_1) f(y_2) \cdots f(y_n), & a < y_1 < y_2 < \cdots < y_n < b \\ 0, & \text{其他} \end{cases} \end{aligned}$$

一种方法是用概率元方法. 即以 $y_1, y_1 + \Delta y_1, ..., y_n, y_n + \Delta y_n$ 为划分, 分成 $2n+1$ 堆的多项分布.

$$\begin{aligned} g(y_1, ..., y_n) =& \lim_{\substack{\Delta y_1 \to 0 \\ \cdots \\ \Delta y_n \to 0}} \frac{P\{y_1 < Y_1 \le y_1 + \Delta y_1, ..., y_n < Y_n \le y_n + \Delta y_n\}}{\Delta y_1 ... \Delta y_n} \\ =& \lim_{\substack{\Delta y_1 \to 0 \\ \cdots \\ \Delta y_n \to 0}} \frac{\frac{n!}{0!1!0!...1!0!}\{F(y_1)\}^0\{F(y_1+\Delta y_1) - F(y_1)\}^1 ...\{F(y_n+\Delta y_n) - F(y_n)\}^1\{1 - F(y_n + \Delta y_n)\}^0}{\Delta y_1 ... \Delta y_n} \\ =& n! \prod_{i=1}^{n} f(y_i). \end{aligned}$$

□

Exercise #12. 13. 令 $(X_1,...,X_n)$ 独立同分布于 $(0,a)$ 的均匀分布. 证明顺序统计量 $(X_{(1)},...,X_{(n)})$ 的联合密度函数为

$$f(y_1,\ldots,y_n)=\begin{cases}\dfrac{n!}{a^n} & \textit{for } y_1<y_2<\ldots<y_n\\ 0 & \textit{otherwise.}\end{cases}$$

证明. 直接应用习题 12.12 的结论即可. □

Exercise #12. 14. 证明若 $X_1,...,X_n$ 是独立同分布于密度函数 f, 分布函数 F, 则 $X_{(k)}$ 的密度函数是

$$f_{(k)}(y)=k\binom{n}{k}f(y)(1-F(y))^{n-k}F(y)^{k-1}.$$

证明. 证明方法一用定义. 用习题 12.12 中的记号, 写出分布函数

$$F_m(x)=P\{Y_k\le x\}=P\left\{\bigcup_{m\ge k}A_m(x)\right\}$$

注意到当 $m\neq n$ 时, $A_m(x)\cap A_n(x)=\emptyset,\forall x\in\mathbb{R}$, 从而

$$P\left\{\bigcup_{m\ge k}A_m(x)\right\}=\sum_{m\ge k}P\{A_m(x)\}$$

而 $P\{A_m(x)\}=\binom{n}{m}\{F(x)\}^m\{1-F(x)\}^{n-m}$ 刚好服从二项分布, 因此

$$F_m(x)=\sum_{m=k}^{n}\binom{n}{m}\{F(x)\}^m\{1-F(x)\}^{n-m}.$$

用归纳法可以证明:

$$\sum_{m=k}^{n}\binom{n}{m}p^m(1-p)^{n-m}=\frac{n!}{(m-1)!(n-m)!}\int_0^p t^{m-1}(1-t)^{n-m}\,dt.$$

于是,

$$F_m(x)=\frac{n!}{(m-1)!(n-m)!}\int_0^{F(x)} t^{m-1}(1-t)^{n-m}\,dt.$$

对 $F_m(x)$ 求导, 即得到 $f_{(k)}(y)$.

另一种证明用概率元方法. 即以 $x,x+\Delta x$ 为划分, 分成三堆的多项分布.

$$\begin{aligned}f_{(k)}(y)&=\lim_{\Delta x\to 0}\frac{P(x<X_{(k)}\le x+\Delta x)}{\Delta x}\\&=\lim_{\Delta x\to 0}\frac{\frac{n!}{(k-1)!1!(n-k)!}\{F(x)\}^{n-1}\{F(x+\Delta x)-F(x)\}^1\{1-F(x+\Delta x)\}^{n-k}}{\Delta x}\\&=\frac{n!}{(k-1)!(n-k)!}\{F(x)\}^{n-1}\lim_{\Delta x\to 0}\frac{\{F(x+\Delta x)-F(x)\}}{\Delta x}\{1-F(x+\Delta x)\}^{n-k}\\&=\frac{n!}{(k-1)!(n-k)!}\{F(x)\}^{n-1}\{1-F(x)\}^{n-k}f(x).\end{aligned}$$

$\square$

Exercise #12. 15 (正态随机变量的模拟)**.** 设 U_1, U_2 是独立的 $(0,1)$ 上的均匀分布. 令 $\theta = 2\pi U_1, S = -\ln(U_2)$.

a) 证明 S 服从指数分布. $R = \sqrt{2S}$ 服从 *Rayleigh* 分布.

b) 令 $X = R\cos(\theta), Y = R\sin(\theta)$. 证明: X, Y 是独立的, 且都服从 $\mu = 0$, $\sigma^2 = 1$ 的正态分布.

注. 这是一种生成正态分布的方法. 被称为 *Box-Muller* 变换.

证明. a) $S = -\ln(U_2)$ 的分布函数为

$$F_S(s) = P\{S \le s\} = P\{-\ln(U_2) \le s\} = P\{U_2 \ge e^{-s}\} = 1 - e^{-s}.$$

从而 S 服从指数分布. $R = \sqrt{2S}$ 的分布函数为

$$F_R(r) = P\{R \le r\} = P\{\sqrt{2S} \le r\} = P\{2S \le r^2\} = P\{S \le \frac{r^2}{2}\} = 1 - e^{-\frac{r^2}{2}}.$$

从而 R 服从 Rayleigh 分布.

b) 由于 θ 和 S 是独立的, 从而 R 和 θ 是独立的. 由于 R 服从参数 $\sigma^2 = 1$ 的 Rayleigh 分布, 从而 X, Y 的联合密度函数为

$$f_{(X,Y)}(x, y) = \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}}.$$

于是 $X = R\cos(\theta), Y = R\sin(\theta)$ 是独立的, X, Y 都服从 $\mu = 0$, $\sigma^2 = 1$ 的正态分布.

$\square$