Chapter 11: 实数上的概率分布

Latest Update: 2025 年 1 月 1 日

Exercise #11. 1. 用 χ^2 随机变量的密度证明 $\Gamma\left(\frac{1}{2}\right)=\sqrt{\pi}$.

证明. 由于 $\chi^2 \sim \Gamma\left(\frac{1}{2},2\right)$, 密度函数为

$$\frac{1}{2^{k/2}\Gamma(k/2)}x^{k/2-1}e^{-x/2}$$

考察 $X\sim N(0,1)$, 则令 $Y=X^2$ 当 $X<0$ 时, $X=-\sqrt{Y}$; 当 $X\geq 0$ 时, $X=\sqrt{Y}$. Jacobian 为 $\frac{1}{2\sqrt{y}}$, 从而 Y 的密度函数为

$$f_Y(y)=2\cdot\frac{1}{2\sqrt{2\pi y}}e^{-y/2}\mathbb{I}_{(0,\infty)}(y)=\frac{1}{\sqrt{2}\Gamma(1/2)}y^{-1/2}e^{-y/2}\mathbb{I}_{(0,\infty)}.$$

于是, $\Gamma\left(\frac{1}{2}\right)=\sqrt{\pi}$. □

Exercise #11. 2. 设 X 是 $[-1,1]$ 上的均匀分布. 求出 $Y=X^k$ 的密度, 其中 k 是正整数.

证明. 由于 X 是 $[-1,1]$ 上的均匀分布, 密度函数为

$$f_X(x)=\frac{1}{2}\mathbb{I}_{[-1,1]}(x).$$

当 k 是奇数时, $g(x)=x^k$ 是一个一一变换, 反函数 $h(y)=g^{-1}(y)=y^{1/k}$, $h'(y)=\frac{1}{k}y^{1/k-1}$, 从而 Y 的密度函数为

$$f_Y(y)=f_X(h(y))\,|h'(y)|=\frac{1}{2k}y^{\frac{1}{k}-1}\mathbb{I}_{[-1,1]}(y).$$

当 k 是偶数时, $g(x)=x^k$ 不是一个一一变换, 但是 $g(x)=x^k$ 分别在 $I_1=[-1,0),I_2=[0,1]$ 上是一个一一变换, 且 $g(I_1)=(0,1],g(I_2)=[0,1]$, 反函数 $h_1(y)=g^{-1}(y)=-y^{1/k}$, $h_2(y)=y^{1/k}$, 且导数为

$$h_1'(y)=-\frac{1}{k}y^{1/k-1},h_2'(y)=\frac{1}{k}y^{1/k-1},$$

从而 Y 的密度函数为

$$f_y(y)=f_X(h_1(y))|h_1'(y)|\mathbb{I}_{(0,1]}(y)+f_X(h_2(y))|h_2'(y)|\mathbb{I}_{[0,1]}(y)=\frac{1}{k}y^{\frac{1}{k}-1}\mathbb{I}_{[0,1]}(y).$$

□

Exercise #11. 3. 设 X 的分布函数是 F. 求出 $Y=|X|$ 的分布函数. 当 X 有一个连续的密度函数 f_X 时, 用 f_X 表达出 Y 的密度函数 f_Y.

证明. $Y=|X|$ 的分布函数为

$$F_Y(y)=P(|X|\le y)=P(-y\le X\le y)=F(y)-F(-y-0).$$

当 X 有一个连续的密度函数 f_X 时,$F_Y(-y-0)=F_Y(-y)$, 于是 $F_Y(y)=F(y)-F(-y)$, 从而 Y 有一个密度函数 f_Y, 且

$$f_Y(y)=F_Y'(y)=F'(y)-F'(-y)=f_X(y)+f_X(-y).$$

□

Exercise #11. 4. 设 X 服从参数为 $(\alpha,1)$ 的 *Cauchy* 分布. 设 $Y=\dfrac{a}{X}$, 其中 $a\neq 0$. 证明: Y 还是一个 *Cauchy* 随机变量, 它的参数是 $\left(\dfrac{a\alpha}{1+\alpha^2},\sqrt{\dfrac{|a|}{1+\alpha^2}}\right)$.

证明. 按教材 P44 关于 Cauchy 分布的定义, Cauchy(α,β) 的 pdf 为

$$f(x)=\frac{1}{\beta\pi}\frac{1}{1+(x-\alpha)^2/\beta^2}.$$

则 X 的密度函数是

$$f_X(x)=\frac{1}{\pi}\frac{1}{1+(x-\alpha)^2}.$$

由于 $Y=g(X)=\dfrac{a}{X}$. 记 $S_0=\{0\},S_1=(0,\infty),S_2=(-\infty,0)$, 从而有 $\mathbb{R}=S_0+S_1+S_2$, 且 g 在 S_1 和 S_2 上是单射, $m_1(S_0)=0$.

在 S_1 上, g 的逆函数为 $g^{-1}(y)=\dfrac{a}{y}$, 从而 g^{-1} 的导数为 $\dfrac{-a}{y^2}$, 在 S_2 上, g 的逆函数为 $g^{-1}(y)=\dfrac{a}{y}$, 从而 g^{-1} 的导数为 $\dfrac{a}{y^2}$. 从而 Y 的密度函数为

$$\begin{aligned}
f_Y(y)=&\frac{1}{\pi}\frac{1}{1+\left(\frac{a}{y}-\alpha\right)^2}\frac{|a|}{y^2}\\
=&\frac{1}{\pi}\frac{|a|}{y^2(1+\alpha^2)+a^2-2a\alpha y}\\
=&\frac{1}{\pi}\frac{|a|}{(1+\alpha^2)\left(y^2-\frac{2a\alpha}{1+\alpha^2}y+\frac{a^2\alpha^2}{(1+\alpha^2)^2}\right)+a^2-\frac{a^2\alpha^2}{1+\alpha^2}}.\\
=&\frac{1}{\pi}\frac{1}{\frac{(1+\alpha^2)}{a}\left(y-\frac{a\alpha}{1+\alpha^2}\right)^2+\frac{|a|}{1+\alpha^2}}\\
=&\frac{1}{\pi\cdot\frac{|a|}{1+\alpha^2}}\frac{1}{\frac{(1+\alpha^2)^2}{a^2}\left(y-\frac{a\alpha}{1+\alpha^2}\right)^2+1}
\end{aligned}$$

从而 Y 服从参数为 $\left(\dfrac{a\alpha}{1+\alpha^2},\dfrac{|a|}{1+\alpha^2}\right)$ 的 Cauchy 分布. □

Exercise #11. 5. 设 X 的密度函数是 f_X, 设 $Y = \dfrac{a}{X}$, 其中 $a \neq 0$. 用 f_X 表示出 Y 的密度函数 f_Y.

证明. 由于 $Y = \dfrac{a}{X}$, Y 的分布函数为

$$F_Y(y) = P(Y \leq y) = P\left(\frac{a}{X} \leq y\right).$$

若 $a > 0, y > 0$, 则

$$F_Y(y) = P\left(X \geq \frac{a}{y}\right) + P(X < 0) = 1 - F_X\left(\frac{a}{y}\right) + F_X(0-0).$$

若 $a > 0, y \leq 0$, 则

$$F_Y(y) = P\left(\frac{a}{y} \leq X < 0\right) = F_X(0-0) - F_X\left(\frac{a}{y} - 0\right).$$

此时, Y 的密度函数为

$$f_Y(y) = F_Y'(y) = f_X\left(\frac{a}{y}\right)\frac{a}{y^2}\mathbb{I}_{(0,\infty)}(y) + f_X\left(\frac{a}{y}\right)\frac{a}{y^2}\mathbb{I}_{(-\infty,0]}(y) = \frac{a}{y^2}f_X\left(\frac{a}{y}\right).$$

若 $a < 0, y > 0$, 则

$$F_Y(y) = P\left(X \leq \frac{a}{y}\right) + P(X > 0) = F_X\left(\frac{a}{y}\right) + 1 - F_X(0-0).$$

若 $a < 0, y \leq 0$, 则

$$F_Y(y) = P\left(0 < X \leq \frac{a}{y}\right) = F_X\left(\frac{a}{y}\right) - F_X(0).$$

此时, Y 的密度函数为

$$f_Y(y) = F_Y'(y) = f_X\left(\frac{a}{y}\right)\frac{-0a}{y^2}\mathbb{I}_{(0,\infty)}(y) + f_X\left(\frac{a}{y}\right)\frac{-a}{y^2}\mathbb{I}_{(-\infty,0]}(y) = \frac{-a}{y^2}f_X\left(\frac{a}{y}\right).$$

综上,

$$f_Y(y) = \frac{|a|}{y^2}f_X\left(\frac{a}{y}\right).$$

□

Exercise #11. 6. 设 X 是 $(-\pi, \pi)$ 上的均匀分布, 令 $Y = \sin(X + \theta)$. 证明: Y 的密度函数是 $f_Y(y) = \dfrac{1}{\pi\sqrt{1-y^2}}$, 其中 $-1 < y < 1$.

证明. 由于 X 是 $(-\pi, \pi)$ 上的均匀分布, 密度函数为

$$f_X(x) = \frac{1}{2\pi}\mathbb{I}_{(-\pi,\pi)}(x).$$

先证明: $P(\sin(X+\theta) \leq y) = P(\sin(X) \leq y), \forall y, \theta$. 事实上, 由于考虑的 $Z = X + \theta$ 取值范围也是一个长度为 2π 的区间, 由于 $\sin$ 是周期函数, 从而积分区域长度一致, 从而 $P(\sin(X+\theta) \leq y) = P(\sin(X) \leq y)$.

接下来, 考察 $Y = \sin(X)$, 令 $g(x) = \sin(x)$, 于是 $g(x)$ 分别在 $I_1 = (-\pi, -\pi/2), I_2 = (-\pi/2, \pi/2), I_3 = (\pi/2, \pi)$ 上是一个一一变换, $I_0 = \{-\pi/2, \pi/2\}$ 是一个零测集, 且 $g(I_1) = (-1, 0), g(I_2) = [-1, 1], g(I_3) = (0, 1)$. 反函数 $h_1(y) = \arcsin(y) - \pi, h_2(y) = \arcsin(y), h_3(y) = \arcsin(y) + \pi$, 导数分别为

$$h_1'(y) = h_2'(y) = h_3'(y) = \frac{1}{\sqrt{1-y^2}}.$$

从而 Y 的密度函数为

$$\begin{aligned}
f_Y(y) &= f_X(h_1(y))|h_1'(y)|\mathbb{I}_{(-1,0)}(y) + f_X(h_2(y))|h_2'(y)|\mathbb{I}_{(-1,1)}(y) + f_X(h_3(y))|h_3'(y)|\mathbb{I}_{(0,1)}(y) \\
&= \frac{1}{2\pi}\frac{1}{\sqrt{1-y^2}}\mathbb{I}_{(-1,0)}(y) + \frac{1}{2\pi}\frac{1}{\sqrt{1-y^2}}\mathbb{I}_{(-1,1)}(y) + \frac{1}{2\pi}\frac{1}{\sqrt{1-y^2}}\mathbb{I}_{(0,1)}(y) \\
&= \frac{1}{\pi\sqrt{1-y^2}}.
\end{aligned}$$

□

Exercise #11. 7. 设 X 有密度函数 f_X. 令 $Y = a\sin(X+\theta)$, 其中 $a > 0$. 证明:

$$f_Y(y) = \frac{1}{\sqrt{a^2-y^2}}\sum_{i=-\infty}^{\infty}\left[f_X\left\{h_i(y)\right\} + f_X\left\{k_i(y)\right\}\right]1_{[-a,a]}(y)$$

对于适当的函数 $h_i(y)$ 和 $k_i(y)$.

证明. 类似上面的讨论, 首先 Y 的分布函数与参数 θ 无关, 从而可以取 $\theta = 0$.

$Y = a\sin(X)$, 令 $g(x) = a\sin(x)$, 则 $g(x)$ 在 $I_k = (-\pi/2 + k\pi, \pi/2 + k\pi)$ 上是一个一一变换, $k \in \mathbb{Z}$, $I_0 = \{\pi/2 + k\pi\}_{k\in\mathbb{Z}}$ 是一个零测集, 且 $g(I_k) = [-a, a]$, $k \in \mathbb{Z}$. 反函数, 当 $k = 2m$ 时, $\phi_k(y) = \arcsin(y/a) + k\pi$, 当 $k = 2m+1$ 时, $\phi_k(y) = \pi - \arcsin(y/a) + k\pi$, $m \in \mathbb{Z}$, 导数分别为

$$\phi_k'(y) = \frac{1}{\sqrt{a^2-y^2}}.$$

记 $h_i(y) = \phi_{2i}(y), k_i(y) = \phi_{2i+1}(y)$, 则 Y 的密度函数为

$$\begin{aligned}
f_Y(y) &= \sum_{m\in\mathbb{Z}}\left[f_X\left\{\phi_{2m}(y)\right\} + f_X\left\{\phi_{2m+1}(y)\right\}\right]\frac{1}{\sqrt{a^2-y^2}}1_{[-a,a]}(y) \\
&= \frac{1}{\sqrt{a^2-y^2}}\sum_{i=-\infty}^{\infty}\left[f_X\left\{h_i(y)\right\} + f_X\left\{k_i(y)\right\}\right]1_{[-a,a]}(y).
\end{aligned}$$

□

Exercise #11. 8. 设 X 是 $(-\pi, \pi)$ 上的均匀分布, 令 $Y = a\tan(X)$, 其中 $a > 0$. 求出 Y 的密度函数 f_Y.

证明. 由于 X 是 $(-\pi, \pi)$ 上的均匀分布, 密度函数为

$$f_X(x) = \frac{1}{2\pi}\mathbb{I}_{(-\pi,\pi)}(x).$$

$g(x)=a\tan(x)$ 分别在 $I_1=(-\pi,-\pi/2), I_2=(-\pi/2,\pi/2), I_3=(\pi/2,\pi)$ 上是一个一一变换, 记 $I_0=\{\pi/2,-\pi/2\}, (-\pi,\pi)=\cup_{i=0}^3 I_i$, 且 I_0 是零测集. 且 $g(I_1)=(0,\infty), g(I_2)=\mathbb{R}^1, g(I_3)=(-\infty,0)$. 于是, 分别单射区域上的反函数 h_i 分别为

$$h_1(y)=\arctan\left(\frac{y}{a}\right)-\pi, h_2(y)=\arctan\left(\frac{y}{a}\right), h_3(y)=\arctan\left(\frac{y}{a}\right)+\pi.$$

导数分别为

$$h_1'(y)=h_2'(y)=h_3'(y)=\frac{1/a}{1+\frac{y^2}{a^2}}=\frac{a}{a^2+y^2}.$$

于是, Y 的密度函数为

$$\begin{aligned}
f_Y(y)&=f_X(h_1(y))|h_1'(y)|\mathbb{I}_{g(I_1)}(y)+f_X(h_2(y))|h_2'(y)|\mathbb{I}_{g(I_2)}(y)+f_X(h_3(y))|h_3'(y)|\mathbb{I}_{g(I_3)}(y)\\
&=\frac{1}{2\pi}\frac{a}{a^2+y^2}\mathbb{I}(y>0)+\frac{1}{2\pi}\frac{a}{a^2+y^2}+\frac{1}{2\pi}\frac{a}{a^2+y^2}\mathbb{I}(y<0)\\
&=\frac{1}{\pi}\frac{a}{a^2+y^2}.
\end{aligned}$$

□

Exercise #11. 9. 设 X 有密度函数, 令

$$Y=ce^{-\alpha X}1_{\{X>0\}}, \quad (\alpha>0, c>0)$$

用 f_X 表示 Y 的密度函数 f_Y.

证明. X 有密度函数, 则 X 的分布函数连续. 由于 $Y=ce^{-\alpha X}1_{\{X>0\}}$, 当 $y>0$ 时, Y 的分布函数为

$$F_Y(y)=P(Y\le y)=P\left(ce^{-\alpha X}1_{\{X>0\}}\le y\right)=P\left(X\ge -\frac{1}{\alpha}\ln\left(\frac{y}{c}\right), X>0\right).$$

当 $y\ge c$ 时, $F_Y(y)=1$; 当 $0<y<c$ 时,

$$F_Y(y)=P\left(X\ge -\frac{1}{\alpha}\ln\left(\frac{y}{c}\right)\right)=1-F_X\left(-\frac{1}{\alpha}\ln\left(\frac{y}{c}\right)\right).$$

此时, Y 的密度函数为

$$f_Y(y)=F_Y'(y)=\frac{1}{\alpha y}f_X\left(-\frac{1}{\alpha}\ln\left(\frac{y}{c}\right)\right)\mathbb{I}_{(0,c)}(y).$$

□

Exercise #11. 10. 称一个密度函数 f 是对称的, 如果对任意 x, 有 $f(x)=f(-x)$. *(即, f 是一个偶函数.)* 若 X 与 $-X$ 是同分布的, 则称随机变量 X 是对称的. 假设 X 的密度函数是 f. 证明: X 是对称的当且仅当 f 是对称的. 在这种情况下, 它是否也允许存在一个非对称的密度函数?

注. 对称密度的例子是: $(-a,a)$ 上的均匀分布; 参数为 $(0,\sigma^2)$ 的正态分布; 参数为 $(0,\beta)$ 的 *Cauchy* 分布, 参数为 $(0,\beta)$ 的双指数分布.

证明. $\Rightarrow$: 若 X 是对称的, 则 X 与 $-X$ 是同分布的,

$$P(X \le x) = P(-X \le x) = P(X \ge -x) = 1 - P(X \le -x).$$

从而 $F(x) = 1 - F(-x)$, 从而 $f(x) = -f(-x)$, 即 f 是一个奇函数.

$\Leftarrow$: 若 f 是对称的, 则 $f(x) = f(-x)$, 从而

$$F(x) = \int_{-\infty}^{x} f(t)dt = \int_{-\infty}^{x} f(-t)dt = \int_{-\infty}^{-x} f(t)dt = 1 - F(-x).$$

从而 X 与 $-X$ 是同分布的.

由于在一个 Lebesgue 零测集上的函数值的改变不会影响积分, 从而可以修改 f 使之不对称, 但是 X 可以对称. $\square$

Exercise #11. 11. 设 X 是正随机变量, 密度函数为 f. 令 $Y = \dfrac{1}{X+1}$, 求出 Y 的密度函数.

证明. 由于 X 是正随机变量, 密度函数为 f. 由于 $Y = \dfrac{1}{X+1} \in (0,1)$, 则 $X = \dfrac{1}{Y} - 1$ 是一个一一变换. 此时, Jacobian $J = \dfrac{\partial X}{\partial Y} = -\dfrac{1}{Y^2} \neq 0$, 从而 Y 的密度函数为

$$\begin{aligned} f_Y(y) &= f\left(\frac{1}{y} - 1\right)\left|-\frac{1}{y^2}\right| \\ &= \frac{1}{y^2} f\left(\frac{1}{y} - 1\right)\mathbb{I}_{(0,1)}(y). \end{aligned}$$

$\square$

Exercise #11. 12. 设 X 是参数为 (μ, σ^2) 的正态分布. 令 $Y = e^X$. 证明: Y 服从对数正态分布.

证明. 若 Y 服从对数正态分布, 则 Y 的密度函数为

$$f(x \mid \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-(\ln x - \mu)^2/2\sigma^2}$$

现在, 考察 $Y = e^X$, 则 $X = \ln Y$. 由于 X 是正态分布, 则 X 的密度函数为

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

从而 Y 的密度函数为

$$\begin{aligned} f_Y(y) &= f_X(\ln y)\left|\frac{1}{y}\right| \\ &= \frac{1}{y\sqrt{2\pi}\sigma} e^{-\frac{(\ln y - \mu)^2}{2\sigma^2}}. \end{aligned}$$

$\square$

Exercise #11. 13. 设 X 是分布函数为 F 的随机变量. 令 $Y = F(X)$. 证明: Y 是 $[0,1]$ 上的均匀分布随机变量.

证明. 定义

$$F^{-1}(u) = \inf\{x : F(x) \geq u\},$$

为 F 的反函数. 考察 Y 的分布函数:

$$P(Y \leq y) = P(F(X) \leq y) = P(X \leq F^{-1}(y)) = F(F^{-1}(y)) = y.$$

□

Exercise #11. 14. 设 F 是一个连续的分布函数, 且 F^{-1} 存在. 设 U 是 $(0,1)$ 上的均匀分布随机变量. 证明: $X = F^{-1}(U)$ 的分布函数是 F.

证明. X 的分布函数为

$$F_X(x) = P(X \leq x) = P(F^{-1}(U) \leq x) = P(U \leq F(x)) = F(x).$$

□

Exercise #11. 15. 设 F 是一个连续的分布函数, U 是 $(0,1)$ 上的均匀分布随机变量. 定义 $G(u) = \inf\{x : F(x) \geq u\}$. 证明: $G(U)$ 的分布函数是 F.

证明. 设 $Y = G(U)$, 则 Y 的分布函数为

$$F(y) = P(Y \leq y) = P(G(U) \leq y) = P(U \leq F(y)) = F(y).$$

得证. □

Exercise #11. 16. 设 $Y = -\dfrac{1}{\lambda}\ln(U)$, 其中 U 是 $(0,1)$ 上的均匀分布随机变量. 证明: Y 服从参数为 λ 的指数分布.

注. 这给出来一种模拟指数随机变量的方法.

证明. 服从参数为 λ 的指数分布的随机变量 X 的密度函数和分布函数分别为:

$$p(x) = \lambda e^{-\lambda x}, x > 0$$
$$F(x) = 1 - e^{-\lambda x}, x > 0$$

根据逆变换法,

$$F^{-1}(x) = -\frac{1}{\lambda}\ln(1-x).$$

从而, $Y = -\dfrac{1}{\lambda}\ln(1-U)$ 服从参数为 λ 的指数分布. 又根据 $U \overset{d}{=} 1-U$, 从而 $Y = -\dfrac{1}{\lambda}\ln(U)$ 也服从参数为 λ 的指数分布. □