连续型随机变量 2 及变换

续本达

清华大学 工程物理系

2023-10-09 清华

复习

随机变量

定义

设 X 是随机变量, 若存在一个非负可积函数 $f(x)$，使得

$$F(x) = \int_{-\infty}^{x} f(t)\mathrm{d}t, -\infty < x < +\infty$$

其中 $F(x)$ 是它的分布函数（别称 累积分布函数 ），则称 X 是 连续型随机变量 ， $f(x)$ 是它的 概率密度函数 。

连续型随机变量

均匀分布

$$f(x) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0, & \text{其他} \end{cases}$$

$$F(x) = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & a \le x < b \\ 1, & x \ge b \end{cases}$$

指数分布

$$f(x;\lambda) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & x \le 0 \end{cases}$$

$$F(x) = \begin{cases} 1-e^{-\lambda x}, & x \ge 0 \\ 0, & x < 0 \end{cases}$$

正态分布

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

$$F(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{x} e^{-\frac{(t-\mu)^2}{2\sigma^2}} \mathrm{d}t$$

伽玛分布

定义

若 X 的概率密度为

$$f(x;\alpha,\lambda)=\frac{\lambda^{\alpha}}{\Gamma(\alpha)}x^{\alpha-1}e^{-\lambda x}, x\geq 0$$

α,λ 为常数，$\alpha>0,\lambda>0$ 则称 X 服从参数为 α,λ 的 伽玛分布 。记作

$$X\sim \mathrm{Ga}(\alpha,\lambda)$$

$$\Gamma(\alpha)=\int_0^{\infty}x^{\alpha-1}e^{-x}\mathrm{d}x$$

为伽玛函数。

定义

若 X 的概率密度为

$$f(x;\alpha,\lambda)=\frac{\lambda^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\lambda x}, x\ge 0$$

α,λ 为常数，$\alpha>0,\lambda>0$ 则称 X 服从参数为 α,λ 的 伽玛分布 。记作

$$X\sim \mathrm{Ga}(\alpha,\lambda)$$

$$\Gamma(\alpha)=\int_0^\infty x^{\alpha-1}e^{-x}\mathrm{d}x$$

为伽玛函数。

特点

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

若 n 是整数，

$$\Gamma(n+1) = n!$$

$$\mathrm{Ga}(1, \lambda) = \mathrm{Exp}(\lambda); X_1, X_2 \sim \mathrm{Exp}(\lambda) \implies X_1 + X_2 \sim \mathrm{Ga}(2, \lambda)$$

即，伽马分布具有可加性。

$$\begin{aligned} f(x; \alpha, \lambda)\mathrm{d}x &= \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} \mathrm{d}x \\ &= \frac{(\lambda x)^{\alpha-1}}{\Gamma(\alpha)} e^{-(\lambda x)} \mathrm{d}(\lambda x) \\ &= f(\lambda x; \alpha, 1)\mathrm{d}(\lambda x) \end{aligned}$$

伽马分布具有伸缩自相似性。

特点

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

若 n 是整数，

$$\Gamma(n+1) = n!$$

$$\mathrm{Ga}(1, \lambda) = \mathrm{Exp}(\lambda); X_1, X_2 \sim \mathrm{Exp}(\lambda) \implies X_1 + X_2 \sim \mathrm{Ga}(2, \lambda)$$

即，伽马分布具有可加性。

$$\begin{aligned} f(x; \alpha, \lambda)\mathrm{d}x &= \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} \mathrm{d}x \\ &= \frac{(\lambda x)^{\alpha-1}}{\Gamma(\alpha)} e^{-(\lambda x)} \mathrm{d}(\lambda x) \\ &= f(\lambda x; \alpha, 1)\mathrm{d}(\lambda x) \end{aligned}$$

伽马分布具有伸缩自相似性。

特点

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

若 n 是整数，

$$\Gamma(n+1) = n!$$

$$\mathrm{Ga}(1,\lambda) = \mathrm{Exp}(\lambda); X_1, X_2 \sim \mathrm{Exp}(\lambda) \implies X_1 + X_2 \sim \mathrm{Ga}(2,\lambda)$$

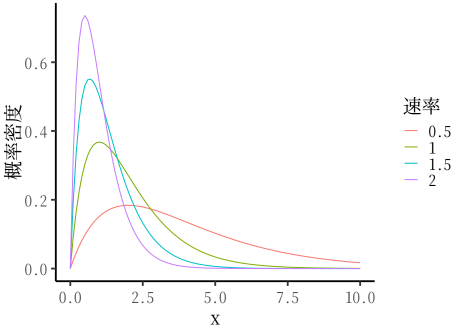
即，伽马分布具有可加性。

$$\begin{aligned} f(x;\alpha,\lambda)\mathrm{d}x =& \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} \mathrm{d}x \\ =& \frac{(\lambda x)^{\alpha-1}}{\Gamma(\alpha)} e^{-(\lambda x)} \mathrm{d}(\lambda x) \\ =& f(\lambda x; \alpha, 1)\mathrm{d}(\lambda x) \end{aligned}$$

伽马分布具有伸缩自相似性。

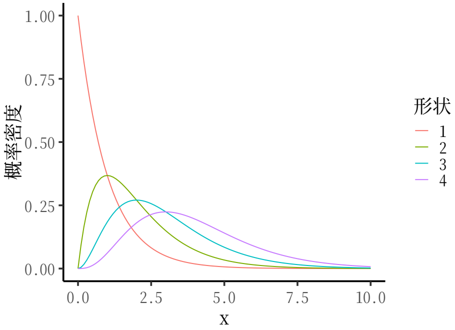
λ 速率参数，设 $\alpha = 2$

```r
xf <- seq(0, 10, by=0.1)
gamma_shape <- data.frame(x=c(), pd=c(), lambda=c())
for (lambda in c(0.5, 1, 1.5, 2)) {
    gamma_shape <- rbind(gamma_shape, data.frame(x=x, pd=dgamma(x, rate=lambda, shape=2), lambda=lambda))
}
gamma_shape$lambda <- as.factor(gamma_shape$lambda)
plot <- ggplot(gamma_shape, aes(x=x, y=pd, color=lambda)) + geom_line()
print(plot + labs(y="概率密度", color="速率"))
```



α 形状参数，设 $\lambda = 1$

```r
x <- seq(0, 10, by=0.1)
gamma_shape <- data.frame(x=c(), pd=c(), alpha=c())
for (alpha in seq(4)) {
    gamma_shape <- rbind(gamma_shape, data.frame(x=x, pd=dgamma(x, scale=1, shape=alpha), alpha=alpha))
}
gamma_shape$alpha <- as.factor(gamma_shape$alpha)
plot <- ggplot(gamma_shape, aes(x=x, y=pd, color=alpha)) + geom_line()
print(plot + labs(y="概率密度", color="形状"))
```



卡方分布

定义

若 n 是整数，$X \sim \mathrm{Ga}\left(\frac{n}{2}, \frac{1}{2}\right)$，称 X 服从 卡方分布 ，其概率密度为

$$f(x;n) = \frac{x^{n/2-1}}{2^{n/2}\Gamma(n/2)} e^{-\frac{x}{2}}, x \geq 0$$

记为

$$X \sim \chi^2(n)$$

在数理统计中有重要地位。

例 $(\chi^2(1))$

$$f(x) = \frac{1}{\sqrt{2\pi x}} e^{-\frac{x}{2}}, x \geq 0$$

- 注意 $\lim_{x\to 0+} f(x) = +\infty$ ，但在 广义积分 的意义下，$F(x)$ 有良定义。

定义

若 n 是整数，$X \sim \mathrm{Ga}\left(\frac{n}{2}, \frac{1}{2}\right)$，称 X 服从 卡方分布 ，其概率密度为

$$f(x;n) = \frac{x^{n/2-1}}{2^{n/2}\Gamma(n/2)} e^{-\frac{x}{2}}, x \geq 0$$

记为

$$X \sim \chi^2(n)$$

在数理统计中有重要地位。

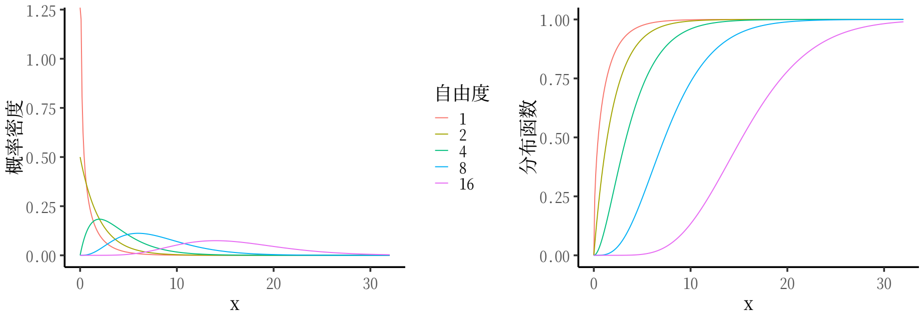
例 $(\chi^2(1))$

$$f(x) = \frac{1}{\sqrt{2\pi x}} e^{-\frac{x}{2}}, x \geq 0$$

- 注意 $\lim_{x\to 0+} f(x) = +\infty$ ，但在 广义积分 的意义下，$F(x)$ 有良定义。

自由度参数 n

```
library(patchwork)
x <- seq(0, 32, by=0.1)
chi_ndf <- data.frame(x=c(), pd=c(), n=c())
for (n in c(1,2,4,8,16)) {
    chi_ndf <- rbind(chi_ndf, data.frame(x=x, pd=dchisq(x, df=n), pp=pchisq(x, df=n), n=n))}
chi_ndf$n <- as.factor(chi_ndf$n)
plot <- ggplot(chi_ndf, aes(x=x, color=n))
d0 <- plot + geom_line(aes(y=pd)) + labs(y="概率密度", color="自由度")
d1 <- plot + geom_line(aes(y=pp), show.legend=FALSE) + labs(y="分布函数")
print(d0 + d1)
```



卡方分布的构造

若 $X_1, X_2, \cdots, X_n \sim N(0,1)$，相互独立，那么

$$X_1^2 + X_2^2 + \cdots + X_n^2 \sim \chi^2(n)$$

证明思路：首先证明 $X_1^2 \sim \chi^2(1)$，见本节课“随机变量的函数分布”例子；随后使用伽马分布的可加性，从 $\chi^2(1)$ 加得 $\chi^2(n)$。

卡方分布的构造

若 $X_1, X_2, \cdots, X_n \sim N(0,1)$ ，相互独立，那么

$$X_1^2 + X_2^2 + \cdots + X_n^2 \sim \chi^2(n)$$

证明思路：首先证明 $X_1^2 \sim \chi^2(1)$ ，见本节课“随机变量的函数分布”例子；随后使用伽马分布的可加性，从 $\chi^2(1)$ 加得 $\chi^2(n)$。

贝塔分布

定义

若 X 的概率密度为

$$f(x;a,b) = \frac{1}{B(a,b)}x^{a-1}(1-x)^{b-1}, 0 < x < 1$$

a, b 为常数，$a > 0, b > 0$，则称 X 服从参数为 a, b 的 贝塔分布 。记作

$$X \sim \mathrm{Be}(a,b)$$

$B(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}\mathrm{d}x$ 为贝塔函数。

对比二项分布

$$\binom{n}{k}p^k(1-p)^{n-k}$$

二项分布 n, p 给定，k 变化。贝塔分布可看成 n, k 给定，p 变化。

定义

若 X 的概率密度为

$$f(x;a,b) = \frac{1}{B(a,b)} x^{a-1}(1-x)^{b-1}, 0 < x < 1$$

a,b 为常数，$a>0, b>0$，则称 X 服从参数为 a,b 的 贝塔分布 。记作

$$X \sim \mathrm{Be}(a,b)$$

$B(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}\mathrm{d}x$ 为贝塔函数。

对比二项分布

$$\binom{n}{k} p^k (1-p)^{n-k}$$

二项分布 n,p 给定，k 变化。贝塔分布可看成 n,k 给定，p 变化。

特点

$$B(a,b) = B(b,a)$$

$$B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

$$\mathrm{Be}(1,1) = U(0,1)$$

$$f(x;a,b)\mathrm{d}x = -f(1-x;b,a)\mathrm{d}x$$

特点

$$B(a,b) = B(b,a)$$

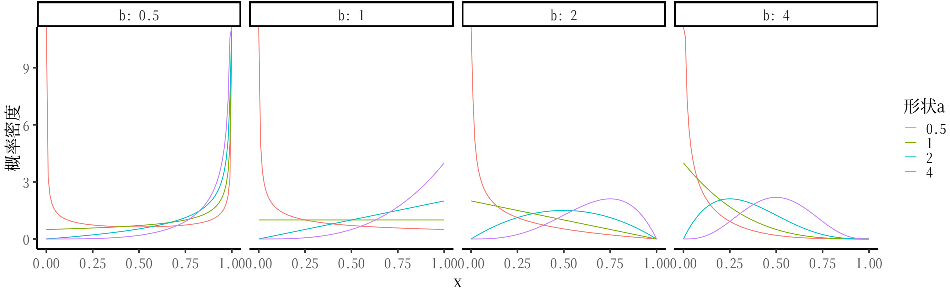
$$B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

$$\mathrm{Be}(1,1) = U(0,1)$$

$$f(x;a,b)\mathrm{d}x = -f(1-x;b,a)\mathrm{d}x$$

$b = 1$ 时，概率密度函数随 a 的变化

```
x <- seq(0, 1, by=0.01)
beta_ab <- data.frame(x=c(), pd=c(), a=c(), b=c())
for (a in c(0.5, 1, 2, 4)) {
    for (b in c(0.5, 1, 2, 4)) {
        beta_ab <- rbind(beta_ab, data.frame(x=x, a=a, b=b,
                                       pd=dbeta(x, shape2=b, shape1=a))) }}
beta_ab$a <- as.factor(beta_ab$a)
beta_ab$b <- as.factor(beta_ab$b)
plot <- ggplot(beta_ab, aes(x=x, y=pd, color=a)) + geom_line()
print(plot + labs(y="概率密度", color="形状 a") + facet_grid(col=vars(b), labeller=label_both))
```



柯西分布

定义

若 X 的概率密度为

$$f(x) = \frac{1}{\pi}\frac{1}{1+x^2}, -\infty < x < \infty$$

则称 X 服从柯西分布。

布莱特-魏格纳分布（Breit-Wigner）

粒子物理中柯西分布常称布莱特-魏格纳分布，描述不稳定粒子的质量的分布，形式为

$$f(x; m_0, \Gamma) = \frac{1}{\pi}\frac{\Gamma/2}{\Gamma^2/4 + (x-m_0)^2}, -\infty < x < \infty$$

- m_0 是粒子的质量，Γ 代表能谱（质量谱）的展宽。

定义

若 X 的概率密度为

$$f(x) = \frac{1}{\pi}\frac{1}{1+x^2}, -\infty < x < \infty$$

则称 X 服从柯西分布。

布莱特-魏格纳分布（Breit-Wigner）

粒子物理中柯西分布常称布莱特-魏格纳分布，描述不稳定粒子的质量的分布，形式为

$$f(x;m_0,\Gamma) = \frac{1}{\pi}\frac{\Gamma/2}{\Gamma^2/4+(x-m_0)^2}, -\infty < x < \infty$$

- m_0 是粒子的质量，Γ 代表能谱（质量谱）的展宽。

朗道分布

定义

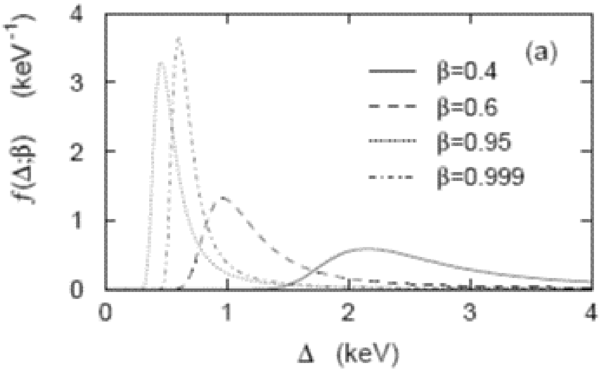
速度为 $\beta = \frac{v}{c}$ 的带电粒子穿过一层厚度为 d 的物质，其能量损失 Δ 服从朗道分布。其概率密度为

$$
\begin{aligned}
f(\Delta;\beta) =& 1/\xi\Phi[\lambda(\Delta)] \\
\Phi(\lambda) =& 1/\pi\int_0^\infty e^{-u(\log u+\lambda)}\sin(\pi u)\mathrm{d}u \\
\lambda(\Delta) =& \frac{1}{\xi}\left[\Delta - \xi\left(\log\frac{\xi}{\epsilon'} + 1 - \frac{1}{\sqrt{1-\beta^2}}\right)\right] \\
\xi =& \frac{2\pi N_A e^4 z^2 \rho(\sum Z)}{m_e c^2(\sum A)}\frac{d}{\beta^2} \\
\epsilon' =& \frac{I^2(1-\beta^2)\exp(\beta^2)}{2m_e c^2\beta^2}
\end{aligned}
$$

I 为平均激发能。

用途

描述粒子的电离能损或能量沉积。



随机变量小结

离散型与连续型随机变量的比较

离散型	连续型
分布律：$p_k = P(X = x_k)$	密度函数：$f(x)\mathrm{d}x$
分布：$F(x) = \sum_{x_i \le x} P(X = x_i)$	分布：$F(x) = \int_{-\infty}^{x} f(t)\mathrm{d}t$
$F(a+) = F(a)$	$P(a < X \le b) = F(b) - F(a)$
点点计算	$P(X = a) = 0$
$F(x)$ 为阶梯函数	$F(x)$ 为连续函数
在 $P(X = a) > 0$ 处 $F(a-) \neq F(a)$	$F(a-) = F(a)$

离散型与连续型随机变量的比较

离散型	连续型
分布律：$p_k = P(X = x_k)$	密度函数：$f(x)\mathrm{d}x$
分布：$F(x) = \sum_{x_i \le x} P(X = x_i)$	分布：$F(x) = \int_{-\infty}^{x} f(t)\mathrm{d}t$
$F(a+) = F(a)$	$P(a < X \le b) = F(b) - F(a)$
点点计算	$P(X = a) = 0$
$F(x)$ 为阶梯函数	$F(x)$ 为连续函数
在 $P(X = a) > 0$ 处 $F(a-) \neq F(a)$	$F(a-) = F(a)$

离散型与连续型随机变量的比较

离散型	连续型
分布律：$p_k = P(X = x_k)$	密度函数：$f(x)\mathrm{d}x$
分布：$F(x) = \sum_{x_i \le x} P(X = x_i)$	分布：$F(x) = \int_{-\infty}^{x} f(t)\mathrm{d}t$
$F(a+) = F(a)$	$P(a < X \le b) = F(b) - F(a)$
点点计算	$P(X = a) = 0$
$F(x)$ 为阶梯函数	$F(x)$ 为连续函数
在 $P(X = a) > 0$ 处 $F(a-) \neq F(a)$	$F(a-) = F(a)$

随机变量的函数分布

引例

我们常对某些随机变量的函数更感兴趣。例如：

- 在一些试验中，所关心的随机变量往往不能由直接测量得到，而它却是某个能直接测量的随机变量的函数。

例 (直径与面积)

我们测量圆轴截面的直径 D，而关心的却是截面面积 $A = \frac{\pi D^2}{4}$。这里，随机变量 A 是随机变量 D 的函数。我们将讨论如何由已知的随机变量 D 的概率分布去求得它的函数 $Y = g(D)$ 的概率分布。

- $g(\cdot)$ 是已知的连续函数

引例

我们常对某些随机变量的函数更感兴趣。例如：

- 在一些试验中，所关心的随机变量往往不能由直接测量得到，而它却是某个能直接测量的随机变量的函数。

例 (直径与面积)

我们测量圆轴截面的直径 D，而关心的却是截面面积 $A = \frac{\pi D^2}{4}$。这里，随机变量 A 是随机变量 D 的函数。我们将讨论如何由已知的随机变量 D 的概率分布去求得它的函数 $Y = g(D)$ 的概率分布。

- $g(\cdot)$ 是已知的连续函数

离散型随机变量的灵感

若 X 是离散型随机变量，其分布律为

X	x_1	x_2	$\cdots$	x_n	$\cdots$
p_k	p_1	p_2	$\cdots$	p_n	$\cdots$

则 $Y=g(x)$ 仍为离散型随机变量，其分布律为

Y	$g(x_1)$	$g(x_2)$	$\cdots$	$g(x_n)$	$\cdots$
p_k	p_1	p_2	$\cdots$	p_n	$\cdots$

如果有 $g(x_1)=g(x_2)$ 怎么办?

注意

$y_i=g(x_i)$ 有相同值时，要合并为一项，对应的概率相加。

离散型随机变量的灵感

若 X 是离散型随机变量，其分布律为

X	x_1	x_2	$\cdots$	x_n	$\cdots$
p_k	p_1	p_2	$\cdots$	p_n	$\cdots$

则 $Y = g(x)$ 仍为离散型随机变量，其分布律为

Y	$g(x_1)$	$g(x_2)$	$\cdots$	$g(x_n)$	$\cdots$
p_k	p_1	p_2	$\cdots$	p_n	$\cdots$

如果有 $g(x_1) = g(x_2)$ 怎么办?

注意

$y_i = g(x_i)$ 有相同值时，要合并为一项，对应的概率相加。

例：离散型随机变量的函数

例

设随机变量 X 具有以下的分布律，试求 $Y=(X-1)^2$ 的分布律。

X	-1	0	1	2
p_k	0.2	0.3	0.1	0.4

Y 所有可能的值为 $0,1,4$

$$\begin{aligned} P(Y=0) &= P(X=1) &&= 0.1 \\ P(Y=1) &= P(X=0)+P(X=2) &&= 0.7 \\ P(Y=4) &= P(X=-1) &&= 0.2 \end{aligned}$$

Y 的分布律为

Y	0	1	4
p_k	0.1	0.7	0.2

例：离散型随机变量的函数

例

设随机变量 X 具有以下的分布律，试求 $Y=(X-1)^2$ 的分布律。

X	-1	0	1	2
p_k	0.2	0.3	0.1	0.4

Y 所有可能的值为 $0,1,4$

$$\begin{aligned} P(Y=0) &= P(X=1) && = 0.1 \\ P(Y=1) &= P(X=0)+P(X=2) && = 0.7 \\ P(Y=4) &= P(X=-1) && = 0.2 \end{aligned}$$

Y 的分布律为

Y	0	1	4
p_k	0.1	0.7	0.2

例：连续型随机变量的函数

例

设随机变量 X 具有概率密度

$$f_X(x) = \begin{cases} \frac{x}{8} & 0 < x < 4, \\ 0 & \text{其他} \end{cases}$$

求随机变量 $Y = 2X + 8$ 的概率密度。

- 记 X, Y 的分布函数分别为 $F_X(x)$ 和 $F_Y(y)$。

$$\begin{aligned} F_Y(y) =& P(Y \leq y) = P(2X + 8 \leq y) \\ =& P\left(X \leq \frac{y}{2} - 4\right) = F_X\left(\frac{y}{2} - 4\right) \end{aligned}$$

例：连续型随机变量的函数

例

设随机变量 X 具有概率密度

$$f_X(x) = \begin{cases} \frac{x}{8} & 0 < x < 4, \\ 0 & \text{其他} \end{cases}$$

求随机变量 $Y = 2X + 8$ 的概率密度。

- 记 X, Y 的分布函数分别为 $F_X(x)$ 和 $F_Y(y)$。

$$\begin{aligned} F_Y(y) =& P(Y \leq y) = P\left(2X + 8 \leq y\right) \\ =& P\left(X \leq \frac{y}{2} - 4\right) = F_X\left(\frac{y}{2} - 4\right) \end{aligned}$$

例：连续型随机变量的函数 – 概率密度

$$f_Y(y) = F'_Y(y) = \frac{\mathrm{d}}{\mathrm{d}y} F_X\left(\frac{y}{2} - 4\right)$$

$$= \frac{1}{2} f_x\left(\frac{y}{2} - 4\right)$$

$$= \begin{cases} \frac{1}{2}\frac{y/2-4}{8} & 0 < \frac{y}{2} - 4 < 4, \\ 0 & \text{其他} \end{cases}$$

$$= \begin{cases} \frac{y-8}{32} & 8 < y < 16, \\ 0 & \text{其他} \end{cases}$$

例：连续型随机变量的函数 – 概率密度

$$
\begin{aligned}
f_Y(y) = F'_Y(y) &= \frac{\mathrm{d}}{\mathrm{d}y} F_X\left(\frac{y}{2} - 4\right) \\
&= \frac{1}{2} f_x\left(\frac{y}{2} - 4\right) \\
&= \begin{cases} \frac{1}{2}\frac{y/2-4}{8} & 0 < \frac{y}{2} - 4 < 4, \\ 0 & \text{其他} \end{cases} \\
&= \begin{cases} \frac{y-8}{32} & 8 < y < 16, \\ 0 & \text{其他} \end{cases}
\end{aligned}
$$

例：连续型随机变量的函数 – 概率密度

$$
\begin{aligned}
f_Y(y) = F_Y'(y) = \frac{\mathrm{d}}{\mathrm{d}y} F_X\left(\frac{y}{2} - 4\right) \\
&= \frac{1}{2} f_x\left(\frac{y}{2} - 4\right) \\
&= \begin{cases} \frac{1}{2}\frac{y/2-4}{8} & 0 < \frac{y}{2} - 4 < 4, \\ 0 & \text{其他} \end{cases} \\
&= \begin{cases} \frac{y-8}{32} & 8 < y < 16, \\ 0 & \text{其他} \end{cases}
\end{aligned}
$$

连续型随机变量函数的概率密度定理

设随机变量 X 的概率密度为 $f_X(x)$,$-\infty < x < \infty$ ，$g(x)$ 处处可导且 $g'(x) > 0$（或恒有 $g'(x) < 0$），则 $Y = g(X)$ 是连续型随机变量，其概率密度为

$$f_Y(y) = \begin{cases} f_X[h(y)]|h'(y)| & a < y < b, \\ 0 & \text{其他} \end{cases}$$

其中 $h(x)$ 为 $g(x)$ 的反函数，$a = \min[g(-\infty), g(\infty)], b = \max[g(-\infty), g(\infty)]$。

微分形式帮助记忆

- 不妨设 $h'(y) > 0$

$$f_X(x)\mathrm{d}x = f_X[h(y)]h'(y)\mathrm{d}y$$
$$\Longrightarrow f_Y(y) = f_X[h(y)]h'(y)$$

连续型随机变量函数的概率密度定理

设随机变量 X 的概率密度为 $f_X(x)$，$-\infty < x < \infty$ ，$g(x)$ 处处可导且 $g'(x) > 0$（或恒有 $g'(x) < 0$），则 $Y = g(X)$ 是连续型随机变量，其概率密度为

$$f_Y(y) = \begin{cases} f_X[h(y)]|h'(y)| & a < y < b, \\ 0 & \text{其他} \end{cases}$$

其中 $h(x)$ 为 $g(x)$ 的反函数，$a = \min[g(-\infty), g(\infty)], b = \max[g(-\infty), g(\infty)]$。

微分形式帮助记忆

- 不妨设 $h'(y) > 0$

$$\begin{aligned} f_X(x)\mathrm{d}x &= f_X[h(y)]h'(y)\mathrm{d}y \\ \Longrightarrow f_Y(y) &= f_X[h(y)]h'(y) \end{aligned}$$

函数的概率密度定理证明

1. 不妨设 $g'(x), h'(y) > 0$。
2. $a = g(-\infty), b = g(\infty)$，则 $y = g(x)$ 的取值范围为 (a, b)。
3. $y < a$ 时， $F_Y(y) = P(Y \leq y) = 0$
4. $y > b$ 时， $F_Y(y) = P(Y \leq y) = 1$
5. $a \leq y \leq b$ 时，

$$\begin{aligned} F_Y(y) =& P(Y \leq y) = P(g(X) \leq y) \\ =& P(X \leq h(y)) = \int_{-\infty}^{h(y)} f_X(x)dx \end{aligned}$$

对 y 求导得

$$f_Y(y) = \begin{cases} f_X[h(y)]|h'(y)| & a < y < b, \\ 0 & \text{其他} \end{cases}$$

函数的概率密度定理证明

1. 不妨设 $g'(x), h'(y) > 0$。
2. $a = g(-\infty), b = g(\infty)$，则 $y = g(x)$ 的取值范围为 (a, b)。
3. $y < a$ 时， $F_Y(y) = P(Y \leq y) = 0$
4. $y > b$ 时， $F_Y(y) = P(Y \leq y) = 1$
5. $a \leq y \leq b$ 时，

$$\begin{aligned} F_Y(y) =& P(Y \leq y) = P(g(X) \leq y) \\ =& P(X \leq h(y)) = \int_{-\infty}^{h(y)} f_X(x)dx \end{aligned}$$

对 y 求导得

$$f_Y(y) = \begin{cases} f_X[h(y)]|h'(y)| & a < y < b, \\ 0 & \text{其他} \end{cases}$$

函数的概率密度定理证明

1. 不妨设 $g'(x), h'(y) > 0$。
2. $a = g(-\infty), b = g(\infty)$，则 $y = g(x)$ 的取值范围为 (a, b)。
3. $y < a$ 时， $F_Y(y) = P(Y \le y) = 0$
4. $y > b$ 时， $F_Y(y) = P(Y \le y) = 1$
5. $a \le y \le b$ 时，

$$\begin{aligned} F_Y(y) =& P(Y \le y) = P(g(X) \le y) \\ =& P(X \le h(y)) = \int_{-\infty}^{h(y)} f_X(x)dx \end{aligned}$$

对 y 求导得

$$f_Y(y) = \begin{cases} f_X[h(y)]|h'(y)| & a < y < b, \\ 0 & \text{其他} \end{cases}$$

例：正态分布伸缩平移

例

设随机变量 $X \sim N(\mu, \sigma^2)$，试证明 X 的线性函数 $Y = aX + b(a \neq 0)$ 也服从正态分布。

- 对于 $-\infty < y < \infty$，有

$$f_Y(y) = f_X[h(y)]|h'(y)| = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{[(y-b)/a-\mu]^2}{2\sigma^2}} \left|\frac{1}{a}\right| = \frac{1}{\sqrt{2\pi}|a|\sigma} e^{-\frac{[y-(a\mu+b)]^2}{2(a\sigma)^2}}$$

因此 $Y \sim N[a\mu + b, (a\sigma)^2]$。

变换

对随机变量 X 的概率密度函数先整体拉伸 a 倍，再平移距离 b。

例：正态分布伸缩平移

例

设随机变量 $X \sim N(\mu, \sigma^2)$，试证明 X 的线性函数 $Y = aX + b(a \neq 0)$ 也服从正态分布。

- 对于 $-\infty < y < \infty$，有

$$f_Y(y) = f_X[h(y)]|h'(y)| = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{[(y-b)/a-\mu]^2}{2\sigma^2}} \left|\frac{1}{a}\right| = \frac{1}{\sqrt{2\pi}|a|\sigma} e^{-\frac{[y-(a\mu+b)]^2}{2(a\sigma)^2}}$$

因此 $Y \sim N[a\mu + b, (a\sigma)^2]$。

变换

对随机变量 X 的概率密度函数先整体拉伸 a 倍，再平移距离 b。

例：柯西分布

例

设随机变量 $X \sim U(-\pi/2, \pi/2)$，求 $Y = \tan X$ 的分布。

- $y = \tan x$ 在 $(-\pi/2, \pi/2)$ 内导函数恒正，取值范围 $(-\infty, \infty)$。
- 逆函数 $x = \tan^{-1} y$ 导函数也恒正，$(\tan^{-1} y)' = \frac{1}{1+y^2}$
- Y 的概率密度为

$$f_Y(y) = f_X[\tan^{-1} y](\tan^{-1} y)' = \frac{1}{\pi}\frac{1}{1+y^2}$$

服从柯西分布。

例：柯西分布

例

设随机变量 $X \sim U(-\pi/2, \pi/2)$，求 $Y = \tan X$ 的分布。

- $y = \tan x$ 在 $(-\pi/2, \pi/2)$ 内导函数恒正，取值范围 $(-\infty, \infty)$。
- 逆函数 $x = \tan^{-1} y$ 导函数也恒正，$(\tan^{-1} y)' = \frac{1}{1+y^2}$
- Y 的概率密度为

$$f_Y(y) = f_X[\tan^{-1} y](\tan^{-1} y)' = \frac{1}{\pi}\frac{1}{1+y^2}$$

服从柯西分布。

例：柯西分布

例

设随机变量 $X \sim U(-\pi/2, \pi/2)$，求 $Y = \tan X$ 的分布。

- $y = \tan x$ 在 $(-\pi/2, \pi/2)$ 内导函数恒正，取值范围 $(-\infty, \infty)$。
- 逆函数 $x = \tan^{-1} y$ 导函数也恒正，$(\tan^{-1} y)' = \frac{1}{1+y^2}$
- Y 的概率密度为

$$f_Y(y) = f_X[\tan^{-1} y](\tan^{-1} y)' = \frac{1}{\pi}\frac{1}{1+y^2}$$

服从柯西分布。

例：反函数不存在的情形

例

设随机变量 X 具有概率密度 $f_X(x), -\infty < x < \infty$，求 $Y = X^2$ 的概率密度。

- 分别记 X, Y 的分布函数为 $F_X(x), F_Y(y)$。
- 由于 $Y = X^2 \geq 0$，所以 $F_Y(y<0) = 0$。当 $y > 0$ 时

$$\begin{aligned} F_Y(y) =& P(Y \leq y) = P(X^2 \leq y) \\ =& P(-\sqrt{y} \leq X \leq \sqrt{y}) = F_X(\sqrt{y}) - F_X(-\sqrt{y}) \end{aligned}$$

- 对 y 求导得

$$f_Y(y) = \begin{cases} \frac{1}{2\sqrt{y}}[f_X(\sqrt{y}) + f_X(-\sqrt{y})] & y > 0 \\ 0 & y \leq 0 \end{cases}$$

例：反函数不存在的情形

例

设随机变量 X 具有概率密度 $f_X(x), -\infty < x < \infty$，求 $Y = X^2$ 的概率密度。

- 分别记 X, Y 的分布函数为 $F_X(x), F_Y(y)$。
- 由于 $Y = X^2 \geq 0$，所以 $F_Y(y < 0) = 0$。当 $y > 0$ 时

$$
\begin{aligned}
F_Y(y) =& P(Y \leq y) = P(X^2 \leq y) \\
=& P(-\sqrt{y} \leq X \leq \sqrt{y}) = F_X(\sqrt{y}) - F_X(-\sqrt{y})
\end{aligned}
$$

- 对 y 求导得

$$
f_Y(y) = \begin{cases} \frac{1}{2\sqrt{y}}[f_X(\sqrt{y}) + f_X(-\sqrt{y})] & y > 0 \\ 0 & y \leq 0 \end{cases}
$$

例：反函数不存在的情形

例

设随机变量 X 具有概率密度 $f_X(x), -\infty < x < \infty$，求 $Y = X^2$ 的概率密度。

- 分别记 X, Y 的分布函数为 $F_X(x), F_Y(y)$。
- 由于 $Y = X^2 \geq 0$，所以 $F_Y(y < 0) = 0$。当 $y > 0$ 时

$$
\begin{aligned}
F_Y(y) =& P(Y \leq y) = P(X^2 \leq y) \\
=& P(-\sqrt{y} \leq X \leq \sqrt{y}) = F_X(\sqrt{y}) - F_X(-\sqrt{y})
\end{aligned}
$$

- 对 y 求导得

$$
f_Y(y) = \begin{cases} \frac{1}{2\sqrt{y}}[f_X(\sqrt{y}) + f_X(-\sqrt{y})] & y > 0 \\ 0 & y \leq 0 \end{cases}
$$

例：反函数不存在的情形

例

设随机变量 X 具有概率密度 $f_X(x), -\infty < x < \infty$，求 $Y = X^2$ 的概率密度。

- 分别记 X, Y 的分布函数为 $F_X(x), F_Y(y)$。
- 由于 $Y = X^2 \geq 0$，所以 $F_Y(y < 0) = 0$。当 $y > 0$ 时

$$
\begin{aligned}
F_Y(y) =& P(Y \leq y) = P(X^2 \leq y) \\
=& P(-\sqrt{y} \leq X \leq \sqrt{y}) = F_X(\sqrt{y}) - F_X(-\sqrt{y})
\end{aligned}
$$

- 对 y 求导得

$$
f_Y(y) = \begin{cases} \frac{1}{2\sqrt{y}}[f_X(\sqrt{y}) + f_X(-\sqrt{y})] & y > 0 \\ 0 & y \leq 0 \end{cases}
$$

从标准正态到卡方分布

若 $X \sim N(0,1)$，则 $y > 0$ 时

$$f_Y(y) = \frac{1}{2\sqrt{y}}\left[\frac{1}{\sqrt{2\pi}}e^{-\frac{y}{2}} + \frac{1}{\sqrt{2\pi}}e^{-\frac{y}{2}}\right] = \frac{1}{\sqrt{2\pi y}}e^{-\frac{y}{2}}$$

是 $n = 1$ 的卡方分布

$$Y \sim \chi^2(1).$$

例：变成均匀分布

例

若随机变量 X 的分布函数 $F_X(x)$ 为严格单调增的连续函数，其反函数 $F_X^{-1}(y)$ 存在，则随机变量 $Y = F_X(X)$ 服从 $(0,1)$ 上的均匀分布，即 $Y \sim U(0,1)$。

1. 由于分布函数 $F_X(x)$ 仅在 $[0,1]$ 内取值，所以 当 $y<0$ 时，$F_Y(y)=0$，当 $y \geq 1$ 时，$F_Y(y)=1$。
2. 当 $0 \leq y < 1$ 时，

$$F_Y(y) = P(Y \leq y) = P(F_X(X) \leq y) = F_X(F_X^{-1}(y)) = y$$

3. 所以 $Y \sim U(0,1)$。

应用

- 可以从 $U(0,1)$ 生成任意连续分布：生成随机数时，只需要实现 $U(0,1)$。

例：变成均匀分布

例

若随机变量 X 的分布函数 $F_X(x)$ 为严格单调增的连续函数，其反函数 $F_X^{-1}(y)$ 存在，则随机变量 $Y \equiv F_X(X)$ 服从 $(0,1)$ 上的均匀分布，即 $Y \sim U(0,1)$。

1. 由于分布函数 $F_X(x)$ 仅在 $[0,1]$ 内取值，所以 当 $y < 0$ 时，$F_Y(y) = 0$，当 $y \geq 1$ 时，$F_Y(y) = 1$。
2. 当 $0 \leq y < 1$ 时，
$$F_Y(y) = P(Y \leq y) = P(F_X(X) \leq y) = F_X(F_X^{-1}(y)) = y$$
3. 所以 $Y \sim U(0,1)$。

应用

- 可以从 $U(0,1)$ 生成任意连续分布：生成随机数时，只需要实现 $U(0,1)$。

例：变成均匀分布

例

若随机变量 X 的分布函数 $F_X(x)$ 为严格单调增的连续函数，其反函数 $F_X^{-1}(y)$ 存在，则随机变量 $Y \equiv F_X(X)$ 服从 $(0,1)$ 上的均匀分布，即 $Y \sim U(0,1)$。

1. 由于分布函数 $F_X(x)$ 仅在 $[0,1]$ 内取值，所以 当 $y < 0$ 时，$F_Y(y) = 0$，当 $y \geq 1$ 时，$F_Y(y) = 1$。
2. 当 $0 \leq y < 1$ 时，

$$F_Y(y) = P(Y \leq y) = P(F_X(X) \leq y) = F_X(F_X^{-1}(y)) = y$$

3. 所以 $Y \sim U(0,1)$。

应用

- 可以从 $U(0,1)$ 生成任意连续分布：生成随机数时，只需要实现 $U(0,1)$。