

求解 $h'(\lambda) = 0$ 可得 $\lambda = \ln(3/2)$. 容易得到在 $(-\infty, \ln(3/2))$ 上导数 $h'(\lambda) > 0$, 而在 $(\ln(3/2), +\infty)$ 上导数 $h'(\lambda) < 0$, 可知 $h(\lambda)$ 在 $\lambda = \ln(3/2)$ 取得最大值. 再根据指数分布的无记忆性有

$$P(X > 5 | X > 3) = P(X > 2) = e^{-2\lambda} = 4/9.$$

4.4.3 正态分布

正态分布是概率统计中最重要的一种分布, 最早由法国数学家棣莫弗 (De Moivre, 1667-1754) 在 1730s 提出, 用于近似抛硬币试验中随机事件的概率, 即中心极限定理的雏形. 德国数学家高斯 (Gauss, 1777-1855) 在 1800s 首次将正态分布应用于预测天文学中星体的位置, 由此才展示出正态分布的应用价值, 后来发现利用正态分布可以描述很多的随机现象, 正态分布也被称为高斯分布.

正态分布在概率统计中的重要性主要体现在以下几方面:

- 现实生活中很多随机现象需要用正态分布进行描述, 如人的身高或体重, 某地区的降雨量等;
- 很多分布可以通过正态分布来进行近似计算, 如后面所学的中心极限定理;
- 数理统计中常用的统计分布是由正态分布导出的, 如后面所学的 χ^2 分布、 t 分布和 F 分布.

定义 4.7 给定任何实数 μ 和 $\sigma > 0$, 若随机变量 X 的概率密度

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

称随机变量 X 服从 **参数为 (μ, σ^2) 的正态分布** (normal distribution), 又称为 **高斯分布** (Gaussian distribution), 记为 $X \sim \mathcal{N}(\mu, \sigma^2)$.

特别地, 当 $\mu = 0$ 和 $\sigma = 1$ 的正态分布 $\mathcal{N}(0, 1)$ 被称为 **标准正态分布**, 此时密度函数为

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad x \in (-\infty, +\infty).$$

对任意 $x \in (-\infty, +\infty)$ 有 $f(x) \geq 0$, 利用极坐标变换 ($x = r \cos \theta, y = r \sin \theta$) 有

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

由此验证了 $\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt = 1$, 在利用简单的变量替换可验证一般正态分布的密度函数.

关于标准正态分布和一般的正态分布, 有如下关系:

定理 4.3 若随机变量 $X \sim \mathcal{N}(\mu, \sigma^2)$, 则有 $Y = (X - \mu)/\sigma \sim \mathcal{N}(0, 1)$; 若随机变量 $X \sim \mathcal{N}(0, 1)$, 则有 $Y = \sigma X + \mu \sim \mathcal{N}(\mu, \sigma^2)$.

证明 若随机变量 $X \sim \mathcal{N}(\mu, \sigma^2)$, 则 $Y = (X - \mu)/\sigma$ 的分布函数

$$F_Y(y) = P[Y \leq y] = P[X - \mu \leq y\sigma] = P[X \leq y\sigma + \mu] = \int_{-\infty}^{\mu+y\sigma} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt.$$

令 $x = (t - \mu)/\sigma$, 代入上面的分布函数有

$$F_Y(y) = \int_{-\infty}^y \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx,$$

由此可证 $Y \sim \mathcal{N}(0, 1)$.

另一方面, 若随机变量 $X \sim \mathcal{N}(0, 1)$, 则有 $Y = \sigma X + \mu$ 的分布函数

$$F_Y(y) = P(Y \leq y) = P(\sigma X + \mu \leq y) = P(X \leq (y - \mu)/\sigma) = \int_{-\infty}^{(y-\mu)/\sigma} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt.$$

令 $t = (x - \mu)/\sigma$, 代入上面的分布函数有

$$F_Y(y) = \int_{-\infty}^y \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx,$$

由此可证 $Y \sim \mathcal{N}(\mu, \sigma^2)$.

关于正态分布的数字特征有

性质 4.10 若随机变量 $X \sim \mathcal{N}(\mu, \sigma^2)$, 则有 $E(X) = \mu$ 和 $\text{Var}(X) = \sigma^2$; 特别地, 若随机变量 $X \sim \mathcal{N}(0, 1)$, 则有 $E(X) = 0$ 和 $\text{Var}(X) = 1$.

正态分布的两个参数分别表示正态分布的期望和方差, 因此数字特征完全刻画了正太分布.

证明 这里仅仅证明标准正态分布的期望为 0 和方差为 1, 结合定理 4.3 可求 $X \sim \mathcal{N}(\mu, \sigma^2)$ 的期望和方差. 若随机变量 $X \sim \mathcal{N}(0, 1)$, 根据奇函数在对称的区间上积分为 0 有

$$E(X) = \int_{-\infty}^{+\infty} \frac{t}{\sqrt{2\pi}} e^{-t^2/2} dt = 0.$$

根据方差的定义和分部积分有

$$\begin{aligned} \text{Var}(X) &= E[X^2] - (E[X])^2 = E[X^2] \\ &= \int_{-\infty}^{+\infty} \frac{t^2}{\sqrt{2\pi}} e^{-t^2/2} dt = -\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} t de^{-t^2/2} \\ &= \left[\frac{te^{-t^2/2}}{-\sqrt{2\pi}} \right]_{t=-\infty}^{+\infty} + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-t^2/2} dt = 1, \end{aligned}$$

由此完成证明.

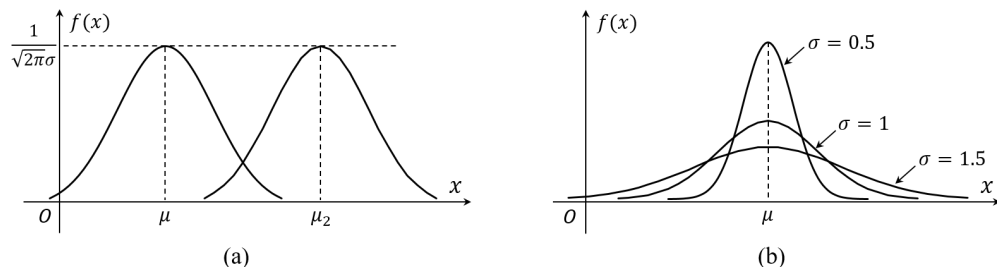


图 4.6 正态分布的密度函数

正态分布的密度函数如图 4.7(a) 所示, 具有以下一些特点:

- 1) 曲线 $f(x)$ 关于 $x = \mu$ 对称, 先单调递增, 之后单调递减, 在 $x = \mu$ 处取最大值 $1/\sqrt{2\pi}\sigma$. 说明随机变量 X 的取值主要集中在 $x = \mu = E(X)$ 附近, 离 $x = \mu$ 越远的区间概率越小.
- 2) 根据 $\lim_{x \rightarrow \pm\infty} f(x) = 0$ 可知曲线 $f(x)$ 以 x 轴为渐近线; 根据 $f''(x) = 0$ 可知曲线 $f(x)$ 的拐点为 $x = \mu \pm \sigma$.
- 3) 固定标准差 σ 而改变期望 μ 的值, 曲线 $f(x)$ 形状不变, 仅沿 x 轴左右平行移动, 如图 4.7(a).
- 4) 固定期望 μ 而改变标准差 σ 的值, 曲线 $f(x)$ 的对称点不变, 但最大值 $1/\sqrt{2\pi}\sigma$ 和拐点 $x = \mu \pm \sigma$ 发生了改变. 如图 4.7(b) 所示: 当 σ 越小, 曲线顶峰越高, 曲线越陡峭, 分布越集中, 方差越小; 反之 σ 越大, 曲线顶峰越低, 曲线越平坦, 分布越分散, 方差越大.

关于正态分布的概率估计, 有下面的不等式:

定理 4.4 若 $X \sim \mathcal{N}(0, 1)$, 对任意 $\epsilon > 0$ 有

$$P(X \geq \epsilon) \leq \frac{1}{2}e^{-\epsilon^2/2}$$

$$P(|X| \geq \epsilon) \leq \min \left\{ 1, \sqrt{\frac{2}{\pi}} \frac{1}{\epsilon} e^{-\epsilon^2/2} \right\}.$$

在上面的定理中, 第一个不等式具有广泛的应用, 在 $\epsilon \in (0, 1)$ 时对真实的概率有更好的估计; 第二个不等式被称为 Mill 不等式, 在 $\epsilon \in (1, +\infty)$ 时对真实的概率有更好的估计.

证明 根据概率密度函数有

$$\begin{aligned}
 P(X \geq \epsilon) &= \int_{\epsilon}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt = \int_0^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-(x+\epsilon)^2/2} dx \\
 &\leq e^{-\epsilon^2/2} \int_0^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = \frac{1}{2} e^{-\epsilon^2/2}.
 \end{aligned}$$

对于 Mill 不等式, 根据 $\mathcal{N}(0, 1)$ 的概率密度 $f(x) = e^{-x^2/2}/\sqrt{2\pi}$ 可得 $f'(x) = -xf(x)$, 于是有

$$\begin{aligned} P(|X| \geq \epsilon) &= 2 \int_{\epsilon}^{+\infty} f(t) dt = 2 \int_{\epsilon}^{+\infty} \frac{tf(t)}{t} dt \\ &\leq 2 \int_{\epsilon}^{+\infty} \frac{tf(t)}{\epsilon} dt = -2 \int_{\epsilon}^{+\infty} \frac{f'(t)}{\epsilon} dt = -\frac{2}{\epsilon} [f(t)]_{\epsilon}^{+\infty} = \frac{2}{\sqrt{2\pi}\epsilon} e^{-\epsilon^2/2}. \end{aligned}$$

由此完成证明.

若随机变量 $X \sim \mathcal{N}(\mu, \sigma^2)$, 则有分布函数

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

该分布函数没有显示的表达式, 只能求数值解, 函数如图 4.7(a) 所示. 为研究正态分布的分布函数, 将其它正态分布转化为标准正态分布 $\mathcal{N}(0, 1)$, 设其分布函数为

$$\Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt.$$

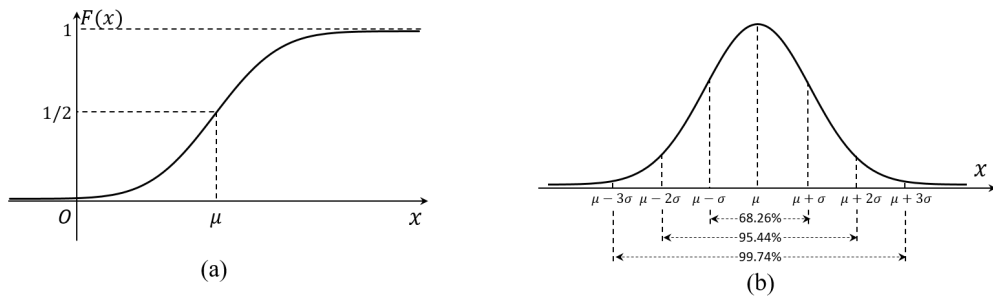


图 4.7 正态分布函数和 3σ 原则

表 4.1 给出了标准正态分布 $\Phi(x)$ 的函数表, 在计算具体的概率时可供查询. 下面给出关于分布函数 $\Phi(x)$ 的一些性质:

- 1) 根据对称性有 $\Phi(x) + \Phi(-x) = 1$.
- 2) 若随机变量 $X \sim \mathcal{N}(\mu, \sigma^2)$, 则对任意实数 $a < b$ 有

$$\begin{aligned} P(X < a) &= P\left(\frac{X-\mu}{\sigma} \leq \frac{a-\mu}{\sigma}\right) = \Phi\left(\frac{a-\mu}{\sigma}\right), \\ P(X > b) &= 1 - P\left(\frac{X-\mu}{\sigma} \leq \frac{b-\mu}{\sigma}\right) = 1 - \Phi\left(\frac{b-\mu}{\sigma}\right), \\ P(a \leq X \leq b) &= P\left(\frac{a-\mu}{\sigma} \leq \frac{X-\mu}{\sigma} \leq \frac{b-\mu}{\sigma}\right) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right). \end{aligned}$$

表 4.1 标准正态分布表 $\Phi(x) = \int_{-\infty}^x e^{-t^2/2}/\sqrt{2\pi}dt$.

x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

3) 若随机变量 $X \sim \mathcal{N}(\mu, \sigma^2)$, 则对任意实数 $k > 0$ 有

$$P(|x - \mu| < k\sigma) = \Phi(k) - \Phi(-k) = 2\Phi(k) - 1.$$

特别的, 当 $k = 1, 2, 3$ 时通过表 4.1 有

$$P(|x - \mu| < \sigma) = 0.6826, \quad P(|x - \mu| < 2\sigma) = 0.9544, \quad P(|x - \mu| < 3\sigma) = 0.9974.$$

如图 4.7(b) 所示, 尽管随机变量 $X \sim \mathcal{N}(\mu, \sigma^2)$ 的取值范围为整个实数域 \mathbb{R} , 但其取值落在 $[\mu - 3\sigma, \mu + 3\sigma]$ 之外的概率不超过千分之三, 也就是 $X \sim \mathcal{N}(\mu, \sigma^2)$ 的取值几乎总在 $[\mu - 3\sigma, \mu + 3\sigma]$ 之内, 这就是人们所说的“ 3σ 原则”, 在实际的统计推断, 特别是产品质量检测中具有重要的应用.

4) 若随机变量 $X \sim \mathcal{N}(\mu, \sigma^2)$, 且已知 $P(X < c) = p$, 则有

$$p = P(X < c) = P\left(\frac{X - \mu}{\sigma} < \frac{c - \mu}{\sigma}\right) = \Phi\left(\frac{c - \mu}{\sigma}\right),$$

由此可反解出 $c = \mu + \sigma\Phi^{-1}(p)$. 这里 $\Phi^{-1}(x)$ 表示标准正态分布函数 $\Phi(x)$ 的反函数, 可根据表 4.1 由里向外查得, 例如 $\Phi^{-1}(0.5871) = 0.22$.

例 4.12 设一公司员工每个月的工资服从正态分布 $\mathcal{N}(6000, \sigma^2)$,

- i) 已知标准差 $\sigma = 500$, 求工资在 5000 与 7000 之间的员工在公司中占比多少?
- ii) 当标准差 σ 为何值时, 工资在 5000 与 7000 之间的员工在公司中占比为 0.803?

解 用随机变量 X 表示公司员工每个月的工资, 则 $X \sim \mathcal{N}(6000, \sigma^2)$. 针对问题 i), 当 $\sigma = 500$ 时通过查询表 4.1 有

$$P(5000 \leq X \leq 7000) = P\left(-2 \leq \frac{X - 6000}{500} \leq 2\right) = \Phi(2) - \Phi(-2) = 2\Phi(2) - 1 = 0.9544.$$

针对问题 ii) 有

$$P(5000 \leq X \leq 7000) = P\left(-\frac{1000}{\sigma} \leq \frac{x - 6000}{\sigma} \leq \frac{1000}{\sigma}\right) = 2\Phi\left(\frac{1000}{\sigma}\right) - 1 = 0.803,$$

于是得到 $\Phi(1000/\sigma) = 0.9015$, 通过由内自外查表 4.1 有 $\sigma \approx 775.2$.

4.5 随机变量函数的分布

已知一个随机变量的概率分布, 可以研究随机变量函数的分布, 例如已知一个圆的半径 X 服从均匀分布 $U(a, b)$, 可以研究圆的面积 $Y = \pi X^2$ 的分布. 针对更一般的情况, 已知随机变量 X 的分布, 如何求解新的随机变量 $Y = g(X)$ 的分布.

设离散型随机变量 X 的分布列为 $p_k = P(X = x_k)$ ($k = 1, 2, \dots$), 求解随机变量 $Y = g(X)$ 的分布列较为简单, 将相等的项 $g(x_i) = g(x_j)$ 合并, 相应的概率相加.

例 4.13 若随机变量 X 的概率分布列为 $P(X = k) = 1/2^k$ ($k = 1, 2, \dots$), 求随机变量 $Y = \cos(\pi X/2)$ 的分布列.

解 当 $n = 1, 2, \dots$ 时有

$$Y = \cos(\pi X/2) = \begin{cases} 1 & X = 4n \\ 0 & X = 2n - 1 \\ -1 & X = 4n - 2, \end{cases}$$

因此 $Y = \cos(\pi X/2)$ 的取值为 $\{-1, 0, +1\}$, 其概率分别为

$$\begin{aligned} P(Y = 1) &= \sum_{n=1}^{\infty} P(X = 4n) = \sum_{n=1}^{\infty} \frac{1}{2^{4n}} = \frac{1}{15}, \\ P(Y = 0) &= \sum_{n=1}^{\infty} P(X = 2n - 1) = \sum_{n=1}^{\infty} \frac{1}{2^{2n-1}} = \frac{2}{3}, \\ P(Y = -1) &= \sum_{n=1}^{\infty} P(X = 4n - 2) = \sum_{n=1}^{\infty} \frac{1}{2^{4n-2}} = \frac{4}{15}, \end{aligned}$$

由此给出随机变量 Y 的分布列.

针对连续型随机变量, 一般采用概率密度函数来刻画其概率分布. 若已知函数 $g(x)$ 和随机变量 X 的概率密度函数为 $f_X(x)$, 求解随机变量 $Y = g(X)$ 的概率密度 $f_Y(y)$, 可分以下两步:

- 1) 求解随机变量 Y 的分布函数 $F_Y(y) = P(Y \leq y) = P(g(X) \leq y) = \int_{g(x) \leq y} f_X(x) dx$;
- 2) 求解随机变量 Y 的密度函数 $f_Y(y) = F'_Y(y)$.

求解此类问题常用到的数学工具是积分求导公式: 设函数 $F(y) = \int_{\psi(y)}^{\varphi(y)} f(x) dx$, 则有

$$F'(y) = f(\varphi(y))\varphi'(y) - f(\psi(y))\psi'(y). \quad (4.2)$$

下面来看两个例子.

例 4.14 设随机变量 X 的概率密度为 $f_X(x)$, 求 $Y = X^2$ 的概率密度.

解 根据分布函数的定义可知

$$F_Y(y) = P(Y \leq y) = P(X^2 \leq y).$$

当 $y \leq 0$ 时有 $F_Y(y) = 0$; 当 $y > 0$ 时有

$$F_Y(y) = P(X^2 \leq y) = P(-\sqrt{y} \leq x \leq \sqrt{y}) = \int_{-\sqrt{y}}^{\sqrt{y}} f_X(x) dx,$$

根据 (4.2) 得到密度函数

$$f_Y(y) = F'_Y(y) = f_X(\sqrt{y})/2\sqrt{y} + f_X(-\sqrt{y})/2\sqrt{y}.$$

综上所述有

$$f_Y(y) = \begin{cases} (f_X(-\sqrt{y}) + f_X(\sqrt{y})) / 2\sqrt{y} & y > 0, \\ 0 & y \leq 0. \end{cases}$$

例 4.15 设随机变量 X 的密度函数为

$$f(x) = \begin{cases} x/8 & 0 < x < 4 \\ 0 & \text{其它,} \end{cases}$$

求 $Y = 2X + 8$ 的密度函数.

解 首先求解分布函数

$$F_Y(y) = P(Y \leq y) = P(2X + 8 \leq y) = P(X \leq (y - 8)/2) = F_X((y - 8)/2),$$

对 y 求导可得 Y 的密度函数

$$f_Y(y) = f_X((y - 8)/2)/2 = \begin{cases} (y - 8)/32 & (y - 8)/2 \in [0, 4] \\ 0 & \text{其它} \end{cases} = \begin{cases} (y - 8)/32 & y \in [8, 16] \\ 0 & \text{其它.} \end{cases}$$

针对一般情况有如下定理:

定理 4.5 设随机变量 X 的概率密度是定义在实数域上的函数 $f_X(x)$, 函数 $y = g(x)$ 处处可导且严格单调 (即 $g'(x) > 0$ 或 $g'(x) < 0$), 令其反函数 $x = g^{-1}(y) = h(y)$, 则 $Y = g(X)$ 的概率密度为

$$f_Y(y) = \begin{cases} f_X(h(y))|h'(y)| & y \in (\alpha, \beta) \\ 0 & \text{其它,} \end{cases}$$

其中 $\alpha = \min\{g(-\infty), g(+\infty)\}$ 和 $\beta = \max\{g(-\infty), g(+\infty)\}$.

可将上述定理推广至区间函数 $x \in [a, b]$, 上述定理依旧成立, 此时有 $\alpha = \min\{g(a), g(b)\}$ 和 $\beta = \max\{g(a), g(b)\}$.

证明 证明思路与前面的例题类似, 这里不妨假设 $g'(x) > 0$ (同理考虑 $g'(x) < 0$). 根据 $g'(x) > 0$ 可知其反函数 $x = h(y)$ 也严格单调, 且 $g(x) \in [\alpha, \beta]$. 因此, 当 $y \leq \alpha$ 时, 有 $F_Y(y) = 0$; 当

$y \geq \beta$ 时有 $F_Y(y) = 1$; 当 $\alpha < y < \beta$ 时,

$$F_Y(y) = P(g(X) < y) = P(X \leq h(y)) = F(h(y)).$$

于是有随机变量 Y 的概率密度

$$f_Y(y) = F'(h(y)) \cdot h'(y) = f_X(h(y)) \cdot h'(y).$$

根据 $x = h(y)$ 严格单调可知 $h'(y) > 0$.

定理 4.6 设 $X \sim \mathcal{N}(\mu, \sigma^2)$, 则 $Y = aX + b$ ($a > 0$) 服从正态分布 $\mathcal{N}(a\mu + b, a^2\sigma^2)$.

证明 设函数 $g(x) = ax + b$, 可得 $\alpha = -\infty$, $\beta = +\infty$, 以及 $y = g(x)$ 的反函数为

$$x = h(y) = (y - b)/a,$$

且有 $h'(y) = 1/a$. 根据定理 4.5 可知

$$f_Y(y) = \frac{1}{a} f_X\left(\frac{y-b}{a}\right) = \frac{1}{a} \frac{1}{\sqrt{2\pi}\sigma} e^{-(\frac{y-b}{a}-\mu)^2/2\sigma^2} = \frac{1}{\sqrt{2\pi}a\sigma} e^{-(y-b-a\mu)^2/2a^2\sigma^2},$$

由此证明了 $Y = aX + b \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$.

例 4.16 设随机变量 X 的分布函数 $F_X(x)$ 是严格单调的连续函数, 则 $Y = F_X(X) \sim U(0, 1)$.

证明 令 $Y = F_X(X)$ 的分布函数为 $F_Y(y)$, 则

$$F_Y(y) = P(Y \leq y) = P(F_X(X) \leq y).$$

因为分布函数 $F_X(x) \in [0, 1]$, 当 $y < 0$ 时有 $F_Y(y) = 0$; 当 $y \geq 1$ 时有 $F_Y(y) = 1$; 当 $y \in [0, 1]$ 时, 由于 $F_X(X)$ 严格单调, 所以 $F_X^{-1}(y)$ 存在且严格单调, 于是有

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

综上所述, 有密度函数

$$f_Y(y) = \begin{cases} 1 & y \in [0, 1] \\ 0 & \text{其它.} \end{cases}$$