第二章　随机变量及其分布

图文并茂详细学习笔记

依据课程 PPT 课件整理

2026 年 11 月 16 日

摘　要

本章引入随机变量，把随机事件的研究转化为变量取值规律的研究，是全书的基石。核心内容：分布函数的定义与性质、离散型随机变量的分布律（0-1 分布、二项分布、泊松分布、超几何分布）、连续型随机变量的概率密度（均匀分布、指数分布、正态分布）以及标准正态化与查表。考试中常以「判定/求分布函数中的常数」「由分布律或密度求概率」「二项分布与泊松近似」「正态分布标准化计算」等形式出现，是必考重点。

目录

本章知识框架 **2**

1　随机变量与分布函数 **2**

1.1　随机变量的概念 2

1.2　分布函数的定义与性质【高频考点】 2

2　离散型随机变量及其分布律 **4**

2.1　分布律及其性质 4

2.2　分布律与分布函数的互求 5

3　常用离散型随机变量的分布【高频考点】 **6**

3.1　两点分布与二项分布 6

3.2　泊松分布与泊松近似【高频考点】 6

3.3　超几何分布 7

4　连续型随机变量及其概率密度 **8**

4.1　概率密度的定义与性质 8

4.2　密度与分布函数互求【高频考点】 8

5　常用连续型随机变量的分布【高频考点】 **9**

5.1　均匀分布 9

5.2　指数分布 9

5.3　正态分布 9

5.4　标准正态分布与标准化【高频考点】 10

5.5　标准正态分布的下侧 α 分位点 11

6　本章总结 **12**

6.1　必背公式汇总 12

6.2　重要求解方法与套路 13

6.3　重要题型归纳 14

6.4　易错点清单 14

本章知识框架

随机变量（样本空间上的实值函数，事件转化为取值）
→ 分布函数 F(x)：定义、三条基本性质、用它表示任意事件的概率
→ 离散型随机变量：分布律及其性质 ↔ 分布函数（阶梯函数）
→ 常用分布：0-1 分布 → 二项分布（n 重贝努里试验）→ 泊松分布（含泊松近似）、超几何分布
→ 连续型随机变量：概率密度 f(x) 及其性质 ↔ 分布函数（连续函数）
→ 常用分布：均匀分布、指数分布、正态分布
→ 标准正态分布：标准化、查表、下侧 α 分位点

1 随机变量与分布函数

1.1 随机变量的概念

定义：设随机试验 E 的样本空间为 $S=\{e\}$，如果对于每一个样本点 $e\in S$，都有确定的实数 $X(e)$ 与之对应，则称此取值为实数的函数 $X=X(e)$ 为**随机变量**，简记为 X。通常用 X,Y,Z 或希腊字母 ξ,η,ζ 等表示。

理解要点（考试常考辨析）：

- 随机变量 X 是定义在样本空间 S 上的**单值实值函数**，它是一个**确定性函数**——不确定性体现在「哪个样本点出现」上，而不是函数本身；
- 引入随机变量后，随机事件可用随机变量的取值来表示，例如「寿命不少于 1000 小时」$=\{X\geq 1000\}$，从而 $P(A)=P\{X\geq 1000\}$；
- 按取值特点，随机变量分为**离散型**（取有限个或可数个值）与**非离散型**（主要研究连续型）。

1.2 分布函数的定义与性质【高频考点】

定义：设 X 为随机变量，对任意实数 x，令

$$F(x)=P\{X\leq x\},\quad -\infty<x<+\infty$$

称 $F(x)$ 为 X 的**（概率）分布函数**，记为 $X\sim F(x)$。

三条基本性质（同时也是「判据」：定义在 $(-\infty,+\infty)$ 上的实函数满足 (1)(2)(3)，则它必是某随机变量的分布函数）：

1. **取值范围**：$0\leq F(x)\leq 1$，且 $F(-\infty)=\lim_{x\to-\infty}F(x)=0$，$F(+\infty)=\lim_{x\to+\infty}F(x)=1$；
2. **单调不减**：$x_1<x_2\Rightarrow F(x_1)\leq F(x_2)$；
3. **右连续**：$F(x+0)=\lim_{t\to x^+}F(t)=F(x)$。

用分布函数表示各种事件的概率（$a<b$）：

- $P\{a<X\leq b\}=F(b)-F(a)$；

- $P\{X=x_0\}=F(x_0)-F(x_0-0)$（$x_0$ 处的跃度）；
- $P\{a\le X\le b\}=F(b)-F(a-0)$；
- $P\{a<X<b\}=F(b-0)-F(a)$；
- $P\{a\le X<b\}=F(b-0)-F(a-0)$。

考场提醒

- 区间左端点开闭决定是否减左极限 $F(a-0)$；右端点开闭决定取 $F(b)$ 还是 $F(b-0)$。口诀：「含端点取函数值，不含端点取左极限」。
- 检验「是否为分布函数」时，最常用来排除的是 $F(+\infty)=1$、$F(-\infty)=0$ 与单调性、右连续。

例 1（分布函数的判定）下列函数中，可作为某随机变量 X 的分布函数的是

(1) $F(x)=1-e^{-x}\ (-\infty<x<+\infty)$；

(2) $F(x)=\frac{1}{2}+\frac{1}{\pi}\arctan x\ (-\infty<x<+\infty)$；

(3) $x>0$ 时 $F(x)=1+e^{-x}$，$x\le 0$ 时 $F(x)=0$；

(4) $|x|<\frac{\pi}{2}$ 时 $F(x)=\frac{1}{2}(1+\cos x)$，$|x|\ge\frac{\pi}{2}$ 时 $F(x)=0$。

解逐条用三性质检验：

- (1)：$x\to-\infty$ 时 $e^{-x}\to+\infty$，$F(x)\to-\infty$，不满足 $0\le F(x)\le 1$，排除；
- (2)：$F(-\infty)=\frac{1}{2}+\frac{1}{\pi}\cdot(-\frac{\pi}{2})=0$，$F(+\infty)=\frac{1}{2}+\frac{1}{\pi}\cdot\frac{\pi}{2}=1$，$\arctan x$ 单调增且连续，满足全部三条性质，**可以作为分布函数**；
- (3)：$x>0$ 时 $F(x)=1+e^{-x}>1$，超出取值范围，排除；
- (4)：$x\ge\frac{\pi}{2}$ 时 $F(x)=0$，但 $F(+\infty)$ 应为 1，排除。

故答案为 **(2)**。

例 2（定常数 + 求概率）已知随机变量 X 的分布函数为：$x>0$ 时 $F(x)=a+be^{-x}$，$x\le 0$ 时 $F(x)=0$。(1) 确定常数 a, b；(2) 求 $P\{X\le\ln 2\}$ 和 $P\{X>1\}$。

解 (1) 由分布函数的性质：

性质 1（$F(+\infty)=1$）：

$$1=F(+\infty)=\lim_{x\to+\infty}(a+be^{-x})=a$$

性质 3（右连续，在 $x=0$ 处）：

$$\lim_{x\to0^+}F(x)=\lim_{x\to0^+}(1+be^{-x})=1+b=F(0)=0$$

所以 $a=1$，$b=-1$，即 $x>0$ 时 $F(x)=1-e^{-x}$，$x\le 0$ 时 $F(x)=0$。

(2) 直接代入分布函数：

$$P\{X\le\ln 2\}=F(\ln 2)=1-e^{-\ln 2}=1-\frac{1}{2}=\frac{1}{2}$$

$$P\{X>1\}=1-P\{X\le 1\}=1-F(1)=1-(1-e^{-1})=e^{-1}$$

思考题：设 $F_1(x),F_2(x)$ 分别为随机变量 X_1,X_2 的分布函数，$a>0,b>0$。若 $F(x)=aF_1(x)+bF_2(x)$ 也是某随机变量的分布函数，则 a, b 必满足什么关系？

解：令 $x\to+\infty$，$F(+\infty)=aF_1(+\infty)+bF_2(+\infty)=a+b=1$，故必有 $a+b=1$（选 D）。

2 离散型随机变量及其分布律

2.1 分布律及其性质

定义：若随机变量 X 只可能取有限个或可数个实数值 $x_1,x_2,\cdots,x_k,\cdots$，则称 X 为**离散型随机变量**。称

$$p_k=P\{X=x_k\},\quad k=1,2,\cdots$$

为 X 的**概率分布**（分布律、分布列）。表示方法有公式法与列表法：

X	x_1	x_2	$\cdots$	x_k	$\cdots$
P	p_1	p_2	$\cdots$	p_k	$\cdots$

两条基本性质（也是判据：满足这两条的一串数必可作为某离散型随机变量的分布律）：

1. $p_k=P\{X=x_k\}\ge 0,\ k=1,2,\cdots$；
2. $\sum_k p_k=1$。

第 2 条是「求分布律中未知常数」的万能钥匙。

例 3（归一性求常数，两小题）

(1) 设 X 的分布律为 $P\{X=k\}=\frac{1-a}{4^k}$，$k=1,2,\cdots$，求 a。

解由 $\sum_{k=1}^{+\infty}P\{X=k\}=1$：

$$\sum_{k=1}^{+\infty}\frac{1-a}{4^k}=(1-a)\sum_{k=1}^{+\infty}\frac{1}{4^k}=(1-a)\cdot\frac{1/4}{1-1/4}=\frac{1-a}{3}=1$$

解得 $a=-2$。

(2) 设 X 的分布律为 $P\{X=k\}=\frac{C}{k!}e^{-2}$，$k=0,1,2,\cdots$，求常数 C。

解

$$\sum_{k=0}^{+\infty}\frac{C}{k!}e^{-2}=Ce^{-2}\sum_{k=0}^{+\infty}\frac{1}{k!}=Ce^{-2}\cdot e=Ce^{-1}=1$$

故 $C=e$。

易错提示

- 求和时注意 k 的起始值：$k=1$ 起与 $k=0$ 起的几何级数/指数级数结果不同；
- 熟记 $\sum_{k=0}^{+\infty}\frac{1}{k!}=e$，$\sum_{k=0}^{+\infty}q^k=\frac{1}{1-q}$（$|q|<1$）。

2.2 分布律与分布函数的互求

定理：设离散型 X 的分布律为 $p_k=P\{X=x_k\}$，$k=1,2,\cdots$，则

1. **由分布律求分布函数**：

$$F(x)=P\{X\le x\}=\sum_{x_k\le x}P\{X=x_k\}=\sum_{x_k\le x}p_k$$

1. **求任意区间概率**：对任意区间 I，$P\{X\in I\}=\sum_{x_k\in I}p_k$；
2. **由分布函数求分布律**：$p_k=P\{X=x_k\}=F(x_k)-F(x_k^-)$（跃度即概率）。

注意

离散型随机变量的分布函数是**阶梯函数**：只有跳跃间断点（在每个取值点 x_k 处），且右连续。写分段区间时一律取「左闭右开」。

例 4（分布律与分布函数互求）将 3 个有区别的球随机地放入木盒和纸盒中，以 X 表示木盒中的球数，求 X 的分布律以及分布函数。

解每个球独立地以概率 $\frac{1}{2}$ 放入木盒，故 $X\sim B(3,\frac{1}{2})$：

$$P\{X=k\}=C_3^k\left(\frac{1}{2}\right)^3,\quad k=0,1,2,3$$

分布律为：

X	**0**	**1**	**2**	**3**
P	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

分段累加求分布函数 $F(x)=P\{X\le x\}$：

- 当 $x<0$：$\{X\le x\}=\emptyset$，$F(x)=0$；
- 当 $0\le x<1$：$F(x)=P\{X=0\}=\frac{1}{8}$；
- 当 $1\le x<2$：$F(x)=P\{X=0\}+P\{X=1\}=\frac{1}{8}+\frac{3}{8}=\frac{1}{2}$；
- 当 $2\le x<3$：$F(x)=\frac{1}{8}+\frac{3}{8}+\frac{3}{8}=\frac{7}{8}$；
- 当 $x\ge 3$：$F(x)=\frac{1}{8}+\frac{3}{8}+\frac{3}{8}+\frac{1}{8}=1$。

x 的范围	$x<0$	$0\le x<1$	$1\le x<2$	$2\le x<3$	$x\ge 3$
$F(x)$	0	$\frac{1}{8}$	$\frac{1}{2}$	$\frac{7}{8}$	1

3 常用离散型随机变量的分布【高频考点】

3.1 两点分布与二项分布

两点分布 (0-1 分布)：若 X 的分布律为 $P\{X=1\}=p$，$P\{X=0\}=q$ $(0<p<1,\ p+q=1)$，则称 X 服从参数为 p 的两点分布。凡只有两个可能结果的试验都可用它描述。

n 重贝努里试验：试验 E 只有两个结果 A 与 $\bar{A}$，$P(A)=p$，$P(\bar{A})=q$ $(p+q=1)$。将 E **独立地**重复 n 次，称为 n 重贝努里试验。有放回抽样、重复掷币、重复射击等都是其模型；无放回抽样在批量很大时也可近似看作贝努里试验。

二项分布：设 X 为 n 重贝努里试验中事件 A 发生的次数，则

$$P\{X=k\}=C_n^k p^k q^{n-k},\quad k=0,1,2,\cdots,n$$

称 X 服从参数为 n, p 的二项分布，记作 $X\sim B(n,p)$。它满足 $P\{X=k\}\geq 0$ 且 $\sum_{k=0}^{n}C_n^k p^k q^{n-k}=(p+q)^n=1$（名称正来源于二项式展开）。两点分布是其特例：$B(1,p)$。

例 5（二项分布直接计算）某射手每次命中目标的概率均为 0.6，现对目标射击 5 次。求：(1) 目标恰好被击中 3 次的概率；(2) 目标被击中的概率。

解把射击一次看作一次试验，射击 5 次即 5 重贝努里试验。设 X 为击中次数，则 $X\sim B(5,0.6)$。

(1) 设 B =「目标恰好被击中 3 次」：

$$P(B)=P\{X=3\}=C_5^3(0.6)^3(0.4)^2=10\times 0.216\times 0.16=0.3456$$

(2) 设 C =「目标被击中」，用对立事件：

$$P(C)=P\{X\geq 1\}=1-P\{X=0\}=1-(0.4)^5=1-0.01024=0.98976$$

考场提醒

「至少发生一次」一律先取对立：$P\{X\geq 1\}=1-P\{X=0\}=1-q^n$。由此也解释了小概率原理：单次概率 ε 再小，$1-(1-\varepsilon)^n\to 1$ $(n\to\infty)$，即小概率事件在大量重复下几乎必然发生。

泊松（Pascal）型分布律举例：一门大炮轰击目标，目标须被击中 r 次才被摧毁，每次命中概率为 p，各次独立，轰击直到摧毁为止。设所需轰击次数为 X，则事件 $\{X=k\}$ 表示「前 $k-1$ 次击中 $r-1$ 次，且第 k 次击中」，故

$$P\{X=k\}=C_{k-1}^{r-1}p^{r-1}(1-p)^{k-r}\cdot p=C_{k-1}^{r-1}p^r(1-p)^{k-r}$$

$k=r,r+1,\cdots$。当 $r=1$ 时即为**几何分布** $P\{X=k\}=p(1-p)^{k-1}$（首次命中所需次数）。

3.2 泊松分布与泊松近似【高频考点】

定义：若 X 的分布律为

$$P\{X=k\}=e^{-\lambda}\cdot\frac{\lambda^k}{k!},\quad k=0,1,2,\cdots$$

其中 $\lambda>0$，则称 X 服从参数为 λ 的**泊松分布**，记作 $X\sim\Pi(\lambda)$。它适于描述**单位时间（或空间）内随机事件发生的次数**：呼叫次数、到达飞机数、布匹疵点数、流星个数等。

泊松定理：设 $\lim_{n\to\infty}np_n=\lambda>0$，则对固定的 k，

$$\lim_{n\to\infty}C_n^kp_n^k(1-p_n)^{n-k}=e^{-\lambda}\frac{\lambda^k}{k!},\quad k=0,1,2,\cdots$$

泊松近似公式：若 $X\sim B(n,p)$，当 n 较大、p 较小而 $np=\lambda$ 适中（经验：$n\geq10$，$p\leq0.1$）时，

$$C_n^kp^k(1-p)^{n-k}\approx e^{-\lambda}\frac{\lambda^k}{k!},\quad \lambda=np$$

例 6（泊松近似 + 查表）某保险公司有 5000 个同年龄的人参加人寿保险：每人年初交付 100 元保险金，一年内死亡则家属领取 3 万元赔偿费。设年死亡率为 1.2‰，求公司这一年至少盈利 20 万元的概率。

解记 X 为 5000 个被保险者中一年内的死亡人数，则 $X\sim B(n,p)$，$n=5000$，$p=0.0012$。设 B =「公司至少盈利 20 万元」。保费总收入 50 万元，盈利 $=50-3X$（万元），故

$$P(B)=P\{50-3X\geq20\}=P\{0\leq X\leq10\}=\sum_{k=0}^{10}C_n^kp^kq^{n-k}$$

n 大 p 小，用泊松近似，$\lambda=np=6$：

$$P(B)\approx\sum_{k=0}^{10}\frac{e^{-6}6^k}{k!}=1-\sum_{k=11}^{+\infty}\frac{e^{-6}6^k}{k!}=1-0.042621\approx0.9574$$

（最后一步查泊松分布表得到。）

同类题套路（课件例，体会列式即可）：

- 次品率 0.016 的集成电路，要求一盒中至少 100 只正品的概率不小于 0.9。设装 $100+N$ 只，次品数 $X\sim B(100+N,0.016)$，$\lambda=(100+N)p\approx1.6$，由 $P\{0\leq X\leq N\}\approx\sum_{k=0}^{N}\frac{e^{-1.6}1.6^k}{k!}\geq0.9$ 查表得 $N\geq3$，至少装 103 只；
- 60 台机床独立工作，故障率 0.02，故障数 $X\sim B(60,0.02)$，$\lambda=np=1.2$。设配 x 名工人，不能及时修理即 $\{X>x\}$，由 $P\{X>x\}\approx\sum_{k=x+1}^{+\infty}\frac{e^{-1.2}1.2^k}{k!}\leq0.01$ 查表得 $x+1\geq5$，即至少配 4 名工人。

3.3 超几何分布

定义：一批产品中有 M 件正品、N 件次品，从中任取 n 件，取到的次品数 X 是离散型随机变量，其分布律为

$$P\{X=k\}=\frac{C_N^kC_M^{n-k}}{C_{M+N}^n},\quad k=0,1,2,\cdots,l,\quad l=\min(n,N)$$

称为**超几何分布**。它对应**无放回抽样**，在产品质量检查与控制中常用；当批量 $M+N$ 远大于抽样数 n 时，可用二项分布 $B(n, \frac{N}{M+N})$ 近似。

易错提示

- 有放回 → 二项分布；无放回 → 超几何分布。审题先看抽样方式；
- 超几何分布中次品取自 N 件次品（C_N^k）、正品取自 M 件正品（C_M^{n-k}），上下标不要配错。

4 连续型随机变量及其概率密度

4.1 概率密度的定义与性质

定义：设随机变量 X 的分布函数为 $F(x)$，如果存在定义在 $(-\infty,+\infty)$ 上的非负可积函数 $f(x)$，使对任何实数 x 恒有

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

则称 X 为**连续型随机变量**，$f(x)$ 称为 X 的**概率密度函数**（概率密度）。

基本性质（判据：满足这两条的可积函数必可作为某连续型随机变量的密度）：

1. 对一切 x，$f(x)\geq 0$（**可以大于 1**）；
2. $\int_{-\infty}^{+\infty} f(x)dx=F(+\infty)=1$。

重要结论：

- $F(x)=\int_{-\infty}^{x} f(t)dt$ 是**连续函数**；
- 若 $f(x)$ 在 x_0 点连续，则 $F(x)$ 在 x_0 可导且 $F'(x_0)=f(x_0)$；若 f 分段连续（有限个间断点），则在连续点处仍有 $F'(x)=f(x)$；
- 连续型随机变量**取任何特定值的概率都是 0**：$P\{X=x_0\}=F(x_0)-F(x_0^-)=0$；
- 因此区间概率与端点开闭无关：$P\{a\leq X\leq b\}=P\{a<X<b\}=F(b)-F(a)=\int_a^b f(x)dx$，几何上即密度曲线下的面积。

思考题：设连续型随机变量 X 的密度为 $f(x)$，分布函数为 $F(x)$，则一定成立的是 ()。A. X 是连续函数；B. $f(x)$ 是连续函数；C. 对任意实数 x 成立 $F'(x)=f(x)$；D. $F(x)$ 是连续函数。

解：选 **D**。f 只需可积、未必连续，故 B、C 不一定成立；随机变量 X 是样本空间上的函数，谈不上是「连续函数」，A 错；而 $F(x)=\int_{-\infty}^{x} f(t)dt$ 必连续。

4.2 密度与分布函数互求【高频考点】

例 7（由密度求概率与分布函数）设随机变量 X 的概率密度为：$-1\leq x<0$ 时 $f(x)=1+x$；$0\leq x<1$ 时 $f(x)=1-x$；其它处 $f(x)=0$。求：(1) $P\{|X|\leq\frac{1}{2}\}$；(2) X 的分布函数。

解 (1)

$$P\left\{|X|\le\frac{1}{2}\right\}=P\left\{-\frac{1}{2}\le X\le\frac{1}{2}\right\}=\int_{-1/2}^{1/2}f(x)dx$$

$$=\int_{-1/2}^{0}(1+x)dx+\int_{0}^{1/2}(1-x)dx=\frac{3}{8}+\frac{3}{8}=\frac{3}{4}$$

(2) 由 $F(x)=\int_{-\infty}^{x}f(t)dt$ 分段讨论：

- 当 $x<-1$：$F(x)=\int_{-\infty}^{x}0\,dt=0$；
- 当 $-1\le x<0$：$F(x)=\int_{-1}^{x}(1+t)dt=\frac{1}{2}x^2+x+\frac{1}{2}$；
- 当 $0\le x<1$：$F(x)=\int_{-1}^{0}(1+t)dt+\int_{0}^{x}(1-t)dt=-\frac{1}{2}x^2+x+\frac{1}{2}$；
- 当 $x\ge1$：$F(x)=\int_{-1}^{0}(1+t)dt+\int_{0}^{1}(1-t)dt=1$。

x 的范围	$x<-1$	$-1\le x<0$	$0\le x<1$	$x\ge1$
$F(x)$	0	$\frac{1}{2}x^2+x+\frac{1}{2}$	$-\frac{1}{2}x^2+x+\frac{1}{2}$	1

思考题（对称密度的概率）：设 X 的密度满足 $f(x)=f(-x)$，则对 $a>0$，$P\{|X|>a\}=$()。A. $1-2F(a)$；B. $2F(a)-1$；C. $2-F(a)$；D. $2[1-F(a)]$。

解：选 **D**。由对称性，

$$P\{|X|>a\}=\int_{-\infty}^{-a}f(t)dt+\int_{a}^{+\infty}f(t)dt=2\int_{a}^{+\infty}f(t)dt=2\left(1-\int_{-\infty}^{a}f(t)dt\right)=2[1-F(a)]$$

5 常用连续型随机变量的分布【高频考点】

5.1 均匀分布

定义：若 X 的概率密度为：$a\le x\le b$ 时 $f(x)=\frac{1}{b-a}$，其它处 $f(x)=0$，则称 X 在区间 $[a,b]$ 上服从**均匀分布**，记作 $X\sim U[a,b]$。

分布函数：$x<a$ 时 $F(x)=0$；$a\le x<b$ 时 $F(x)=\frac{x-a}{b-a}$；$x\ge b$ 时 $F(x)=1$。

要点：X 落在 $[a,b]$ 内任一子区间的概率只与区间长度成正比（几何概率的连续版）。

5.2 指数分布

定义：若 X 的概率密度为：$x\ge0$ 时 $f(x)=\lambda e^{-\lambda x}$，$x<0$ 时 $f(x)=0$（$\lambda>0$ 为常数），则称 X 服从参数为 λ 的**指数分布**，记作 $X\sim e(\lambda)$。

分布函数（由 $F(x)=\int_{-\infty}^{x}f(t)dt$ 推得）：$x<0$ 时 $F(x)=0$；$x\ge0$ 时 $F(x)=1-e^{-\lambda x}$。

要点：指数分布是各种「寿命」的近似模型（元件寿命、通话时间、服务时间等），在可靠性理论中地位重要。常用 $P\{X>t\}=e^{-\lambda t}$（$t>0$）。

5.3 正态分布

概率积分（推导正态密度归一性的基础，记住结论）：

$$\int_{-\infty}^{+\infty} e^{-x^2}dx = \sqrt{\pi}$$

定义：若连续型随机变量 X 的概率密度为

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}, \quad -\infty < x < +\infty$$

其中 $-\infty < \mu < +\infty$，$\sigma > 0$ 均为常数，则称 X 服从参数为 μ, σ 的**正态分布**，记作 $X \sim N(\mu, \sigma^2)$。换元 $y = \frac{x-\mu}{\sqrt{2}\sigma}$ 配合概率积分可验证 $\int_{-\infty}^{+\infty} f(x)dx = 1$。

密度曲线性质：

- 关于直线 $x = \mu$ 对称，$(-\infty, \mu]$ 上严格增，$[\mu, +\infty)$ 上严格减；
- $x = \mu$ 处取最大值 $f(\mu) = \frac{1}{\sigma\sqrt{2\pi}}$；
- $\lim_{x\to\pm\infty} f(x) = 0$，曲线以 x 轴为渐近线；
- 曲线在 $x = \mu \pm \sigma$ 处有拐点。

5.4 标准正态分布与标准化【高频考点】

定义：$\mu = 0, \sigma = 1$ 的正态分布 $N(0,1)$ 称为**标准正态分布**，其密度与分布函数专用记号：

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2}\right\}, \quad \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left\{-\frac{t^2}{2}\right\} dt$$

**$\Phi(x)$ 的性质 **：

1. $\Phi(x) + \Phi(-x) = 1$；
2. $\Phi(0) = \frac{1}{2}$；
3. $\Phi'(x) = \varphi(x) > 0$，即 $\Phi(x)$ 严格单调递增；
4. $\Phi(x)$ 可查标准正态分布表。

标准化公式（一般正态 → 标准正态，一切正态概率计算的总钥匙）：设 $X \sim N(\mu, \sigma^2)$，则

$$F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right), \quad P\{a < X \le b\} = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

常用查表值：$\Phi(1) = 0.8413$，$\Phi(1.645) = 0.95$，$\Phi(1.96) = 0.975$，$\Phi(2) = 0.9772$，$\Phi(3) = 0.9987$。

例 8（对称区间概率与 3σ 法则）设 $X \sim N(\mu, \sigma^2)$，$\lambda > 0$，证明 X 落在 $(\mu - \lambda\sigma, \mu + \lambda\sigma)$ 内的概率只与 λ 有关，与 μ, σ 无关。

解

$$P\{\mu - \lambda\sigma < X < \mu + \lambda\sigma\} = F(\mu + \lambda\sigma) - F(\mu - \lambda\sigma)$$

$$= \Phi\left(\frac{\mu + \lambda\sigma - \mu}{\sigma}\right) - \Phi\left(\frac{\mu - \lambda\sigma - \mu}{\sigma}\right) = \Phi(\lambda) - \Phi(-\lambda) = 2\Phi(\lambda) - 1$$

结果只含 λ，证毕。当 $\lambda=3$ 时 $P\{\mu-3\sigma<X<\mu+3\sigma\}=2\Phi(3)-1=99.74\%$，这就是 **3$\sigma$ 法则**（正态变量几乎总落在均值的 3 倍标准差之内）。

例 9（反用标准化求参数信息）已知 $X\sim N(2,\sigma^2)$，且 $P\{|X-3|\le 1\}=0.44$，求 $P\{|X-2|\ge 2\}$。

解先由已知条件定出 $\Phi(\frac{2}{\sigma})$：

$$P\{|X-3|\le 1\}=P\{2\le X\le 4\}=F(4)-F(2)=\Phi\left(\frac{2}{\sigma}\right)-\Phi(0)=\Phi\left(\frac{2}{\sigma}\right)-0.5=0.44$$

故 $\Phi(\frac{2}{\sigma})=0.94$。再求所求概率（注意 $|X-2|\ge 2$ 即 $X\le 0$ 或 $X\ge 4$）：

$$P\{|X-2|\ge 2\}=1-P\{|X-2|<2\}=1-P\{0<X<4\}$$

$$=1-[F(4)-F(0)]=1-\left[\Phi\left(\frac{2}{\sigma}\right)-\Phi\left(-\frac{2}{\sigma}\right)\right]=2\left[1-\Phi\left(\frac{2}{\sigma}\right)\right]=2\times 0.06=0.12$$

思考题（混合密度）：设 X 的分布函数为 $F(x)$，密度 $f(x)=af_1(x)+bf_2(x)$，其中 $f_1(x)$ 是正态分布 $N(0,\sigma^2)$ 的密度，$f_2(x)$ 是参数为 λ 的指数分布的密度，且 $a>0,b>0$。已知 $F(0)=\frac{1}{8}$，求 a, b。

解：由 $f=af_1+bf_2$ 知 $F(x)=aF_1(x)+bF_2(x)$ 且 $a+b=1$。又 $F_1(0)=\Phi(0)=\frac{1}{2}$，指数分布在 $x=0$ 处 $F_2(0)=0$，故

$$F(0)=a\cdot\frac{1}{2}+b\cdot 0=\frac{a}{2}=\frac{1}{8}\Rightarrow a=\frac{1}{4},\ b=\frac{3}{4}$$

5.5 标准正态分布的下侧 α 分位点

定义：设 $X\sim N(0,1)$，给定 α（$0<\alpha<1$），若存在唯一 z_α 使

$$P\{X\le z_\alpha\}=\Phi(z_\alpha)=\alpha$$

则称 z_α 为 $N(0,1)$ 的**下侧 α 分位点**（分位数）。

性质（$0<\alpha<1$）：

1. $P\{X>z_{1-\alpha}\}=\alpha$；
2. $z_{1-\alpha}=-z_\alpha$（对称性）；
3. $P\{|X|>z_{1-\alpha/2}\}=\alpha$，即 $P\{-z_{1-\alpha/2}\le X\le z_{1-\alpha/2}\}=1-\alpha$。

例 10（用分位点表示常数）设 $X\sim N(\mu,\sigma^2)$，试用分位点表示常数 a, b：(1) $\mu=0,\sigma=1$，$P\{-X<a\}=0.025$；(2) $\mu=1,\sigma=2$，$P\{|X-1|\le b\}=0.75$。

解 (1) $X\sim N(0,1)$，

$$P\{-X<a\}=P\{X>-a\}=1-P\{X\le -a\}=0.025\Rightarrow P\{X\le -a\}=0.975$$

故 $-a=z_{0.975}$，即 $a=-z_{0.975}=z_{0.025}$。

(2) 标准化：$Y=\frac{X-1}{2}\sim N(0,1)$，

$$P\{|X-1|\le b\}=P\left\{|Y|\le\frac{b}{2}\right\}=2\Phi\left(\frac{b}{2}\right)-1=0.75\Rightarrow\Phi\left(\frac{b}{2}\right)=0.875$$

故 $\frac{b}{2}=z_{0.875}$，即 $b=2z_{0.875}$。

例 11（正态 + 二项综合）设测量某目标距离的随机误差 X 具有概率密度

$$f(x)=\frac{1}{40\sqrt{2\pi}}\exp\left\{-\frac{(x-20)^2}{3200}\right\},\quad -\infty<x<+\infty$$

即 $X\sim N(20,40^2)$。求四次独立测量中，至少有一次误差的绝对值不超过 20m 的概率。

解先求单次测量「误差绝对值不超过 20m」的概率：

$$p=P\{|X|\le 20\}=F(20)-F(-20)=\Phi\left(\frac{20-20}{40}\right)-\Phi\left(\frac{-20-20}{40}\right)=\Phi(0)-\Phi(-1)$$

$$=0.5-[1-\Phi(1)]=0.8413-0.5=0.3413$$

令 Y 为四次测量中误差绝对值不超过 20m 的次数，则 $Y\sim B(4,p)$。设 C =「至少有一次误差的绝对值不超过 20m」，则

$$P(C)=P\{Y\ge 1\}=1-P\{Y=0\}=1-(1-p)^4=1-(1-0.3413)^4\approx 0.8117$$

同类题：某仪器装有 4 只独立工作的同类元件，每只寿命 $X\sim N(5000,\sigma^2)$（小时），工作元件不少于 2 只时仪器正常工作。则由对称性 $p=P\{X>5000\}=1-\Phi(0)=\frac{1}{2}$，能工作 5000 小时以上的元件数 $Y\sim B(4,\frac{1}{2})$，仪器正常工作 5000 小时以上的概率为

$$P\{Y\ge 2\}=1-P\{Y=0\}-P\{Y=1\}=1-C_4^0\left(\frac{1}{2}\right)^4-C_4^1\left(\frac{1}{2}\right)^4=1-\frac{1}{16}-\frac{4}{16}=\frac{11}{16}$$

6 本章总结

6.1 必背公式汇总

1. 分布函数定义：$F(x)=P\{X\le x\}$；
2. 分布函数三性质：$0\le F(x)\le 1$ 且 $F(-\infty)=0$，$F(+\infty)=1$；单调不减；右连续 $F(x+0)=F(x)$；
3. 事件概率：$P\{a<X\le b\}=F(b)-F(a)$；$P\{X=x_0\}=F(x_0)-F(x_0-0)$；
4. 分布律两性质：$p_k\ge 0$，$\sum_k p_k=1$；且 $F(x)=\sum_{x_k\le x}p_k$，$p_k=F(x_k)-F(x_k^-)$；
5. 两点分布：$P\{X=1\}=p$，$P\{X=0\}=q$（$p+q=1$）；

6. 二项分布 $X\sim B(n,p)$：$P\{X=k\}=C_n^k p^k q^{n-k}$，$k=0,1,\cdots,n$；
7. 泊松分布 $X\sim \Pi(\lambda)$：$P\{X=k\}=e^{-\lambda}\frac{\lambda^k}{k!}$，$k=0,1,2,\cdots$；
8. 泊松近似：n 大 p 小时 $C_n^k p^k(1-p)^{n-k}\approx e^{-\lambda}\frac{\lambda^k}{k!}$，$\lambda=np$（$n\geq 10$，$p\leq 0.1$）；
9. 超几何分布：$P\{X=k\}=\frac{C_N^k C_M^{n-k}}{C_{M+N}^n}$，$k=0,1,\cdots,\min(n,N)$；
10. 密度两性质：$f(x)\geq 0$，$\int_{-\infty}^{+\infty}f(x)dx=1$；$F(x)=\int_{-\infty}^{x}f(t)dt$，$F'(x)=f(x)$（连续点处）；
11. 连续型区间概率：$P\{a\leq X\leq b\}=F(b)-F(a)=\int_a^b f(x)dx$，且 $P\{X=x_0\}=0$；
12. 均匀分布 $U[a,b]$：$f(x)=\frac{1}{b-a}$（$a\leq x\leq b$），$F(x)=\frac{x-a}{b-a}$（$a\leq x<b$）；
13. 指数分布 $e(\lambda)$：$x\geq 0$ 时 $f(x)=\lambda e^{-\lambda x}$，$F(x)=1-e^{-\lambda x}$，$P\{X>t\}=e^{-\lambda t}$；
14. 正态分布 $N(\mu,\sigma^2)$：$f(x)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}$，最大值 $f(\mu)=\frac{1}{\sigma\sqrt{2\pi}}$，拐点 $x=\mu\pm\sigma$；
15. 标准化：$F(x)=\Phi\left(\frac{x-\mu}{\sigma}\right)$；$\Phi(-x)=1-\Phi(x)$，$\Phi(0)=\frac{1}{2}$；
16. 对称区间与 3σ 法则：$P\{|X-\mu|<\lambda\sigma\}=2\Phi(\lambda)-1$，$\lambda=3$ 时为 99.74%；
17. 下侧 α 分位点：$\Phi(z_\alpha)=\alpha$，$z_{1-\alpha}=-z_\alpha$；
18. 概率积分：$\int_{-\infty}^{+\infty}e^{-x^2}dx=\sqrt{\pi}$。

6.2 重要求解方法与套路

套路一：判定/构造分布函数（或密度、分布律）第 1 步：写下对应的三条（或两条）基本性质；第 2 步：优先代入极限条件（$F(+\infty)=1$、$F(-\infty)=0$ 或 $\sum p_k=1$、$\int f=1$）定常数；第 3 步：验证单调性/非负性/右连续。

** 套路二：已知 $F(x)$ 求概率 ** 第 1 步：把事件改写为 $\{X\leq\cdot\}$ 或区间形式；第 2 步：套 $P\{a<X\leq b\}=F(b)-F(a)$，端点开闭用左极限调整；第 3 步：单点概率 $P\{X=x_0\}=F(x_0)-F(x_0-0)$。

套路三：分布律 $\leftrightarrow$ 分布函数（离散型）第 1 步：列出全部取值与概率；第 2 步：按 $F(x)=\sum_{x_k\leq x}p_k$ 分段累加，区间一律左闭右开；第 3 步：反向时用跃度 $p_k=F(x_k)-F(x_k^-)$。

套路四：密度 $\leftrightarrow$ 分布函数（连续型）第 1 步：按密度的分段点划分区间；第 2 步：每段上 $F(x)=\int_{-\infty}^{x}f(t)dt$ 逐段累加积分；第 3 步：求概率直接对密度在目标区间积分，不必管端点开闭。

套路五：正态分布求概率第 1 步：标准化，$F(x)=\Phi(\frac{x-\mu}{\sigma})$；第 2 步：用 $\Phi(-x)=1-\Phi(x)$ 化为正变量；第 3 步：查表求值；若与「至少一次」复合，先算单次概率 p，再用 $1-(1-p)^n$（$Y\sim B(n,p)$）。

6.3 重要题型归纳

题型	识别特征	解题套路/所用公式
判断是否为分布函数	给若干函数问哪个可作 $F(x)$	验三性质，先用 $F(\pm\infty)$ 与值域排除
分布函数/分布律/密度定常数	表达式含待定常数	$F(+\infty)=1$、右连续；$\sum p_k=1$；$\int f=1$
由 $F(x)$ 求事件概率	已知 $F(x)$ 求 $P\{a<X\le b\}$ 等	$F(b)-F(a)$，注意端点开闭与左极限
离散型求分布律与 $F(x)$	古典概型背景（放球、取产品）	先求各 p_k（常为二项），再分段累加
二项分布应用	n 次独立重复试验，求恰发生 k 次/至少一次	$C_n^k p^k q^{n-k}$；至少一次用 $1-q^n$
泊松近似计算	n 大（≥10）p 小（≤0.1）的二项问题	$\lambda=np$，换 $e^{-\lambda}\lambda^k/k!$ 再查表
由密度求概率与 $F(x)$	分段密度	分段积分 $F(x)=\int_{-\infty}^{x} f(t)dt$
正态分布计算	$X\sim N(\mu,\sigma^2)$ 求概率/反求参数	标准化 $F(x)=\Phi(\frac{x-\mu}{\sigma})$，查表
分位点表示	用 z_α 表示未知常数	化为 $\Phi(\cdot)=\alpha$ 或对称区间形式
正态与二项综合	「n 次独立测量/元件至少…」	先正态求单次 p，再 $Y\sim B(n,p)$

6.4 易错点清单

- 分布函数是右连续，离散型的分段区间必须写成左闭右开，写反即错；
- 求 $P\{a\le X\le b\}$（含端点）时忘记把 $F(a)$ 换成 $F(a-0)$；
- 把 $P\{X=x_0\}$ 当成 0——只有连续型才是 0，离散型单点概率等于跃度；
- 密度 $f(x)$ 可以大于 1（如 $U[0,\frac{1}{2}]$），「密度 ≤ 1」是错误直觉；
- $F'(x)=f(x)$ 只在 f 的连续点成立，不能无条件下结论；
- 二项分布与超几何分布混用：有放回用二项，无放回用超几何；
- 泊松近似忘记 $\lambda=np$，或查表时把 $\sum_{k=0}^{N}$ 与 $\sum_{k=N+1}^{+\infty}$（用 1 减去）弄混；
- 正态标准化时把 σ^2 当 σ 用，如 $N(20,40^2)$ 应除以 40 而非 1600；
- 忘用 $\Phi(-x)=1-\Phi(x)$ 导致负值查不了表；
- 分位点定义看清是「下侧」：$\Phi(z_\alpha)=\alpha$，且 $z_{1-\alpha}=-z_\alpha$；
- 连续型「$X\sim e(\lambda)$」是课件记号（指数分布），勿与常数 e 混淆；
- 「至少一次」不取对立事件直接求和，计算量翻倍且易错。