2.10 高斯分布

2.10.1 高斯分布的定义

如果连续随机变量X的概率密度函数为：

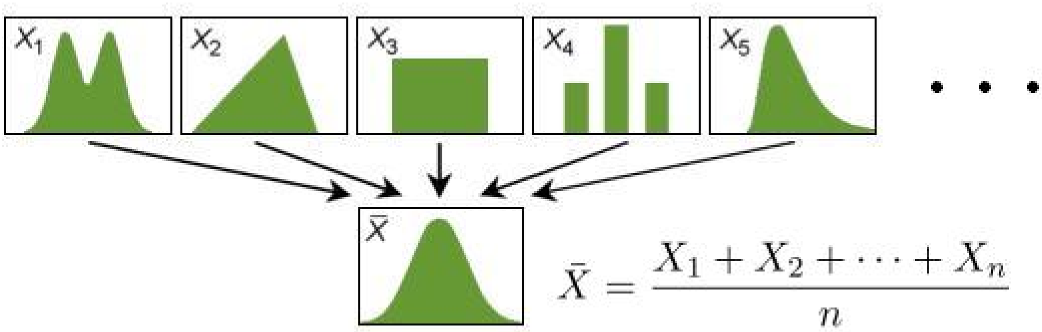
$$p(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}},\quad -\infty<x<+\infty$$

则称X服从正态分布（normal distribution），也称作高斯分布（Gaussian distribution），记作$X\sim N(\mu,\sigma^2)$，其累积分布函数为：

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

2.10.2 中心极限定理

因为要逼近二项分布才引入了高斯分布，但随着这个问题研究的深入，发现不光二项分布，很多别的分布最终也会变为高斯分布，甚至很多不同的分布混合在一起最终也会变为高斯分布：



高斯分布是这些分布的最终归宿，这是一个很惊人的结论，这就导致高斯分布的地位大大提高，基本可以算是概率论的中心。

这个结论称为中心极限定理，它导致的结果是，我们在现实生活中会观察到非常非常多的高斯分布。

2.10.3 标准高斯分布

我们称$\mu=0$、$\sigma=1$时的高斯分布$N(0,1)$为标准正态分布。

通常记标准正态变量为Z（Z Score），记标准高斯分布的概率密度函数为$\phi(x)$，累积分布函数为$\Phi(x)$，即：

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

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

2.10.4 转为标准高斯分布

2.10.4.1 转化

高斯分布有一个重要结论：

如果$X \sim N(\mu, \sigma^2)$，那么有：

$$aX + b \sim N(a\mu + b, a^2\sigma^2)$$

通过这个结论，我们知道，只需要设置合适a、b就可以让X转为标准高斯分布，具体而言就是令（在概率与统计中，Z基本上就是标准高斯分布随机变量的专用符号）：

$$Z = \frac{X - \mu}{\sigma}$$

则有（请自行验证）：

$$Z \sim N(0, 1)$$

> 例题
>
> 如果X服从正态分布，参数$\mu = 3$和$\sigma^2 = 9$,求$P(2 < X < 5)$
>
> A
>
> $$\Phi\left(\frac{2}{3}\right) - \Phi\left(\frac{1}{3}\right)$$
>
> B正确
>
> $$\Phi\left(\frac{2}{3}\right) - \Phi\left(-\frac{1}{3}\right)$$
>
> 根据正态变量的标准化，将正态分布转换成标准正态分布
>
> $$P(2 < X < 5) = P\left(\frac{2-3}{3} < \frac{X-3}{3} < \frac{5-3}{3}\right)$$
>
> $$= P\left(-\frac{1}{3} < Z < \frac{2}{3}\right)$$
>
> $$= \Phi\left(\frac{2}{3}\right) - \Phi\left(-\frac{1}{3}\right)$$
>
> **2**
>
> 假设从你家到办公室索要花费的事件服从均值为40分钟，标准差为7分钟的正态分布。
>
> 如果你要参加一个下午1点的会议，地点就在办公室，那么想要做到95%的把握不迟到，那么最晚何时出发
>
> 设X是路上花的事件，我们需要确定x，使

> P(X>x)=1-0.95=0.05
>
> 因为服从正态分布，所以它等价于
>
> $$P\left(\frac{X-40}{7}>\frac{x-40}{7}\right)=0.05$$
>
> 即我们需要求满足
>
> $$P\left(Z>\frac{x-40}{7}\right)=0.05$$
>
> 的x，其中Z为标准正态随机变量
>
> 根据正态分布常用公式得
>
> $$P\left(Z>\frac{x-40}{7}\right)=1-\Phi\left(\frac{x-40}{7}\right)$$
>
> 再根据标准正态分布表得
>
> $$P(Z>1.65)<0.05<P(Z>1.64)$$
>
> 因此
>
> $$1.64<\frac{x-40}{7}<1.65$$
>
> 即
>
> $$51.48<x<51.55$$
>
> 这样，你应该在12点过8.45分与12点过8.52分之间动身。

2.10.4.2 上α分位点

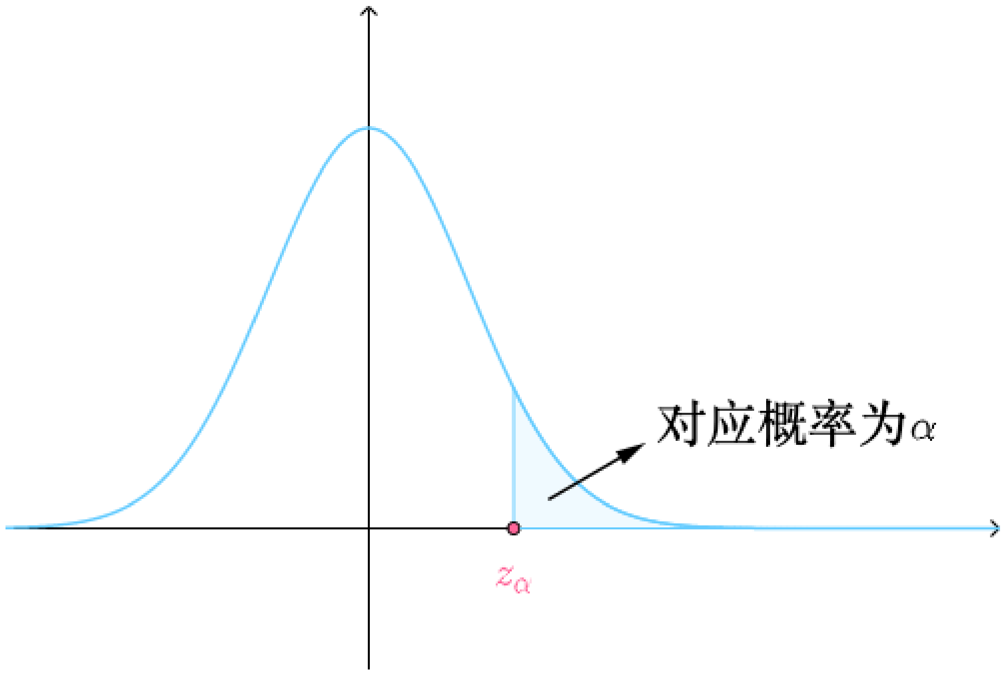
因为高斯分布都可以转为标准高斯分布，所以对于标准高斯分布数学家们有更多的研究和定义：

如果有$Z\sim N(0,1)$，如果z_α满足：

$$P(Z>z_\alpha)=\alpha,\quad 0<\alpha<1$$

那么称点z_α为标准高斯分布的上α分位点。

从几何上看，z_α就是满足下面要求的点：



2.10.5 期望与方差

高斯分布$X \sim N(\mu, \sigma^2)$的期望和方差为：

$$E(X) = \mu, \quad Var(X) = \sigma^2$$

2.11 指数分布

2.11.1 指数分布的推导与定义

馒头店每段时间内卖出的馒头数量服从泊松分布，而两次卖出馒头之间的时间间隔大于t的概率，根据分析，等同于t时间内没有卖出一个馒头的概率，而后者的概率可以由泊松过程给出。至此所需的条件都齐备了，那么开始解题吧，假设随机变量：

$$Y = \text{两次卖出馒头之间的时间间隔}$$

这个随机变量的概率可以如下计算：

$$P(Y > t) = P(X = 0, t) = \frac{(\lambda t)^0}{0!} e^{-\lambda t} = e^{-\lambda t}, \quad t \geq 0$$

进而有：

$$P(Y \leq t) = 1 - P(Y > t) = 1 - e^{-\lambda t}$$

这其实已经得到了Y的累积分布函数了，所以我们可以得到一个概率分布：

定义：

若随机变量X的概率密度函数为：

$$p(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

其中$\lambda > 0$，称X服从指数分布，也可以记为：

$$X \sim Exp(\lambda)$$

累积分布函数为：

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

由以上可得，指数分布可以用于描述相邻随机事件发生的时间间隔，即可以理解为连续的几何分布。

> 例题：
>
> **1**
>
> 某站台上，两辆公交车到站的时间间隔是否服从指数分布？
>
> 这个问题其实是问，在一段时间内到达站台的公交车数量是否满足泊松分布的三个条件：
>
> - **平稳性**：在一段时间内公交车到达的数量还是差不多的（可能需要分时段统计，比如白天和夜间，高峰与低谷）
> - **普通性**：可以认为公交车进站总有先后，不会在同一瞬间到达两辆公交车
> - **独立性**：公交车到达车站的事件彼此独立
>
> 因此符合泊松分布，进而两辆公交车之间的时间间隔也符合指数分布。
>
> **2**
>
> 电气设备的寿命是否服从指数分布？
>
> 假想电器的内部有一枚硬币，抛出反面的概率很大，而正面很不容易抛出。每一秒抛一次硬币，如果是反面，那么电器还可以使用，如果是正面，那么电器就寿终正寝。那么它的寿命就可以看作是二项分布（或者说几何分布，抛到一次正面就结束），由于抛硬币的次数很多（即n很大），所以最终可以看作是泊松分布。那么它的寿命相当于时间间隔，确实会服从指数分布。

2.11.2 指数分布的期望与方差

指数分布$X \sim Exp(\lambda)$的期望和方差为：

$$E(X)=\frac{1}{\lambda}, \quad Var(X)=\frac{1}{\lambda^2}$$

先计算：

$$E(X)=\int_0^{+\infty} x\lambda e^{-\lambda x}\mathrm{d}x$$

$$=\int_0^{+\infty} x\mathrm{d}\left(-e^{-\lambda x}\right)$$

$$=-xe^{-\lambda x}\Big|_0^{+\infty}+\int_0^{+\infty} e^{-\lambda x}\mathrm{d}x=\frac{1}{\lambda}$$

又有：

$$
\begin{aligned}
E(X^2) &= \int_0^{+\infty} x^2 \lambda e^{-\lambda x} \mathrm{d}x \\
&= \int_0^{+\infty} x^2 \mathrm{d}\left(-e^{-\lambda x}\right) \\
&= -x^2 e^{-\lambda x}\Big|_0^{+\infty} + 2\int_0^{+\infty} x e^{-\lambda x} \mathrm{d}x = \frac{2}{\lambda^2}
\end{aligned}
$$

所以:

$$
Var(X) = E(X^2) - \mu^2 = \frac{1}{\lambda^2}
$$

2.11.3 指数分布的无记忆性

如果已经有一天没有卖出馒头了，那么明天卖不出馒头的概率会不会减少?

把这个问题抽象下，假设卖出馒头的时间间隔X服从指数分布:

$$
X \sim Exp(\lambda)
$$

已经s天没有卖出馒头的概率为:

$$
P(X > s) = e^{-\lambda s}
$$

在这个条件下，t天后还是卖不出馒头的概率为:

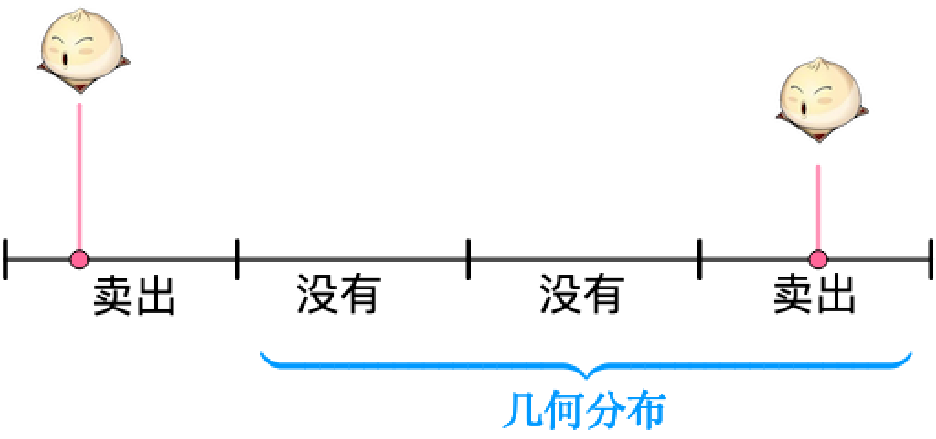
$$
\begin{aligned}
P(X > s + t | X > s) &= \frac{P(X > s + t)}{P(X > s)} = \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} \\
&= e^{-\lambda t} = P(X > t)
\end{aligned}
$$

就好像之前s天没有卖出馒头这个事情不存在一样，这就叫作指数分布的无记忆性。

2.11.4 指数分布与几何分布

几何分布大概描述的是抛硬币，扔了无数次反面之后终于得到正面:

卖馒头也可以看作扔硬币，没有卖出相当于得到反面，卖出相当于得到正面，那么两次卖馒头之间相当于有一个几何分布:



所以指数分布和几何分布内在都是相同的，因此都具有无记忆性；这两者也是唯二具有无记忆性的分布。

2.12 一维分布总结

离散随机变量的概率分布：

伯努利分布	抛硬币，二选一
二项分布	n重伯努利，出现k次“是”
泊松分布	二项分布的极限
几何分布	n重伯努利，第k次首次出现“是”
负二项分布	几何分布的和
超几何分布	不放回抽样的二项分布

连续随机变量的概率分布：

均匀分布	古典派中的几何概型
正态分布	二项分布的另外一种极限
指数分布	泊松分布的间隔，连续的几何分布