目录

文件名

第一部分　正态分布 and 标准正态分布

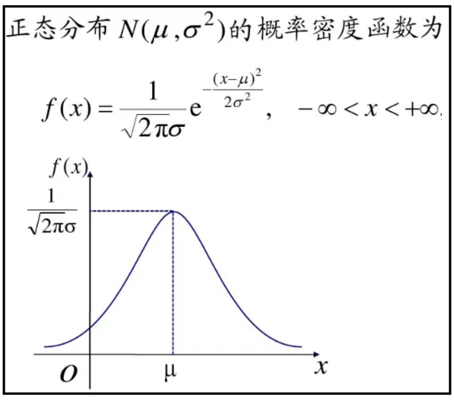
1　正态分布, normal distribution

正态分布, normal distribution, 直译过来就是 “最常态下的分布”, “一般最常见的分布”.
正态分布, 是概率分布中最重要的分布. 在数学家眼里, 它是远远高于其他分布的. 有很多其他的分布，比如对数正态分布、T 分布、F 分布, 都是直接由 “正态分布” 推导出来的.

1.1　正态分布 - 概率密度函数: $\boxed{\varphi(x)=\frac{1}{\sqrt{2\pi}\cdot\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}}$

概率密度函数. 用小写的 φ 表示.
“正态分布” $N(\underbrace{\mu}_{\text{平均值.}},\underbrace{\sigma^2}_{\sigma\text{是标准差. }\sigma^2\text{是方差}})$ 的概率函数是: $\boxed{\varphi(x)=\frac{1}{\sqrt{2\pi}\cdot\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}\ (-\infty<x<+\infty)}$

记作: $\boxed{X\sim N(\mu,\sigma^2)}$ ← 称为: X 服从 “参数为 μ, σ 的正态分布 (或高斯分布)”.
- 这里的 N, 就是正态分布 (Normal distribution) 的英文首字母.
- μ 是 “平均值”
- σ 是 “标准差”
- 注意: 概率函数公式里, 这第二个参数写的是 σ^2, 而不是 σ! 所以, 比如对于 N(1, 100) 来说, 其 $\mu=1$, $\sigma^2=100$, 即 $\sigma=10$.



1.2　正态分布 - 累加函数: $\boxed{F(x)=\frac{1}{\sqrt{2\pi}\cdot\sigma}\int_{-\infty}^{x}\left[e^{-\frac{(x-\mu)^2}{2\sigma^2}}\right]dx}$

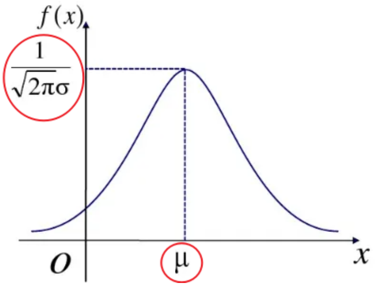
对 “概率函数 f(x)” 求积分, 其曲线下的阴影面积就是 “累加函数 F(x)”. 其面积 =1.

$$
\begin{aligned}
\underbrace{F(x)}_{\text{累加函数}} &= \int \underbrace{f(x)}_{\text{概率函数}} \, dx \\
&= \int_{-\infty}^{x} \left[\frac{1}{\sqrt{2\pi} \cdot \sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \right] dx \\
&= \frac{1}{\sqrt{2\pi} \cdot \sigma} \int_{-\infty}^{x} \left[e^{-\frac{(x-\mu)^2}{2\sigma^2}} \right] dx
\end{aligned}
$$

1.3 正态分布 (钟形曲线) 的 "概率函数" 的性质

1.3.1 以 "x= 均值 μ" 为对称轴

正态分布的 "概率函数" 曲线, 以 "x= 均值 μ" 为对称轴. 在此处, 函数的 y 值达到最大. 即此时 $y = \dfrac{1}{\sqrt{2\pi} \cdot \sigma}$

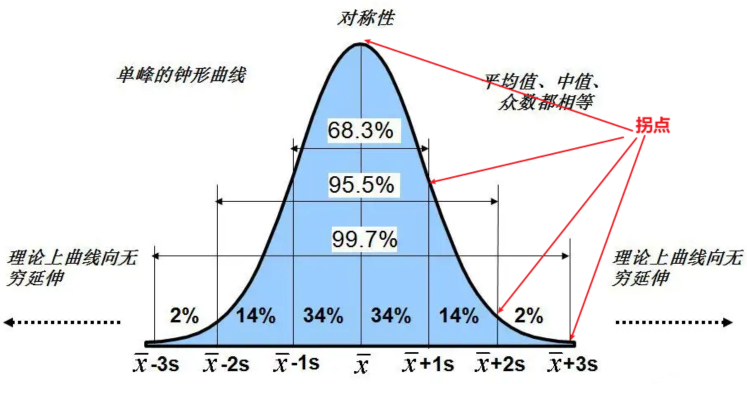


所以, 对称轴 μ, 能控制图像的 "左右平移".

1.3.2 以 x 轴为渐近线

就是说, 曲线的两端, 无限接近于 y=0, 而不会掉落到 -y 领域上去.

1.3.3 在 $x = \mu \pm \sigma$ 处有 "拐点"



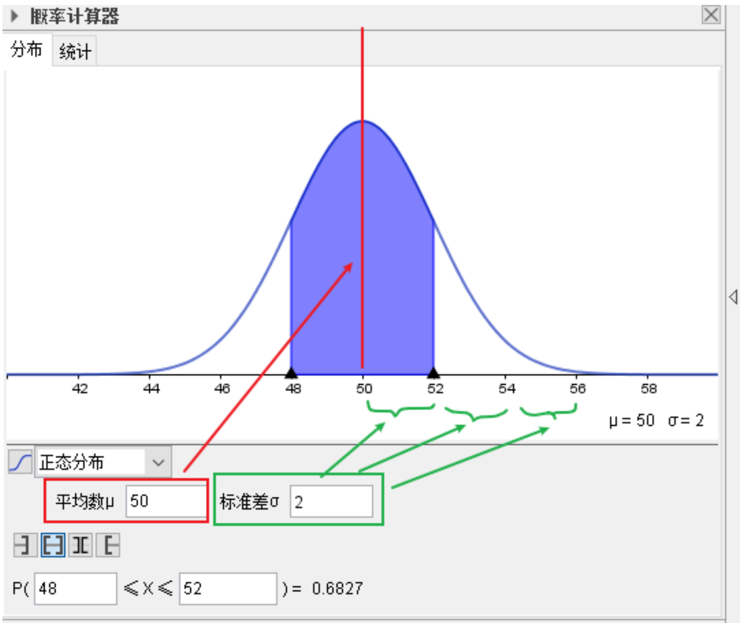
正态分布

1.3.4 "σ 标准差" 参数, 控制图像的 "矮胖" 或 "高瘦"

→ 若 σ 变小: 因为在 x= 处, y 有最大值是 $\dfrac{1}{\sqrt{2} \cdot \sigma}$. 所以当 σ 变小时, 分母变小, 则分数值就变大, 即 y 值变大, 所以图像会拉高, 变瘦高.

→ 若 σ 变大: 则最高点的 y 值变小, 图像会压低, 变矮胖.

但注意, 无论是变瘦高, 还是变矮胖, 曲线下的阴影面积, 始终是 =1, 不变的!



---

2 标准正态分布 ← 即当均值 $\mu=0$, 标准差 $\sigma=1$ 时的 "正态分布"

2.1 标准正态分布 - 概率函数: $\boxed{\phi_0(x)=f(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}}$

我们把 $\mu=0,\sigma=1$, 代入正态分布的 PDF 和 CDF 函数中, 就得到:

"标准正态分布" 的 "概率函数 PDF" (专门记作 $\phi_0(x)$) :

$$\phi_0(x)=f(x)=\frac{1}{\sqrt{2\pi}\cdot\underbrace{\sigma}_{=1}}e^{-\frac{(x-\overbrace{\mu}^{=0})^2}{2\sigma^2}}=\frac{1}{\sqrt{2\pi}\cdot 1}e^{-\frac{(x-0)^2}{2\cdot 1^2}}=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$$

即: $\boxed{\phi_0(x)=f(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}}$

2.2 标准正态分布 - 累加函数: $\boxed{\Phi_0(x)=F(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}[e^{-\frac{x}{2}}]dx}$

"标准正态分布" 的 "累加函数 CDF" (专门记作 $\Phi_0(x)$) :

$$\begin{aligned}\Phi_0(x)&=F(x)\\&=\frac{1}{\sqrt{2\pi}\cdot\underbrace{\sigma}_{=1}}\int_{-\infty}^{x}\left[e^{-\frac{(x-\overbrace{\mu}^{=0})^2}{2\sigma^2}}\right]dx\\&=\frac{1}{\sqrt{2\pi}\cdot 1}\int_{-\infty}^{x}\left[e^{-\frac{(x-0)^2}{2\cdot 1^2}}\right]dx\\&=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}[e^{-\frac{x}{2}}]dx\end{aligned}$$

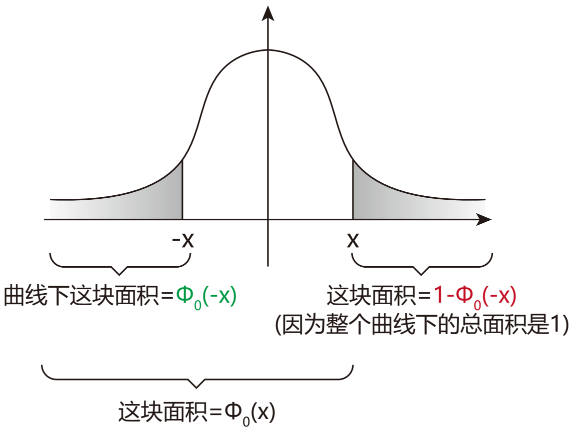
即: $\boxed{\Phi_0(x)=F(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\left[e^{-\frac{x}{2}}\right]dx}$

2.3 标准正态分布的性质

2.3.1 因为它的 $\mu=0$, 所以它的函数曲线, 关于 x=0 对称, 即 y 轴是对称轴.

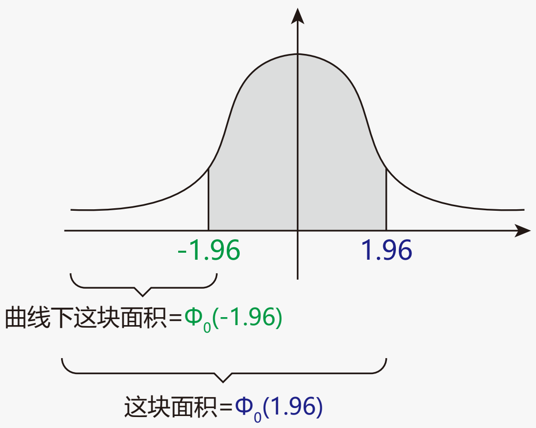
所以它就是个偶函数. 有:
- 概率函数 $\varphi_0(x)=\varphi_0(-x)$ ← 我们在下标处加个 0, 来表示它是 "标准" 的正态分布函数的 "概率函数" 或 "累加函数".

- 其 "累加函数" 有: $\boxed{\Phi_0(-x)=1-\Phi_0(x)}$ ← 这个公式很重要!
比如: $\Phi_0(-4)=1-\Phi_0(4)$



例

$$
\begin{aligned}
\text{求 } P\{|x|\leq 1.96\} &= P\{-1.96\leq X\leq 1.96\}\\
&=\Phi_0(1.96)-\underbrace{\Phi_0(-1.96)}_{=1-\Phi_0(1.96)}\\
&=\Phi_0(1.96)-1+\Phi_0(1.96)\\
&=2\Phi_0(1.96)-1
\end{aligned}
$$



两者相减, 就是中间阴影部分的面积

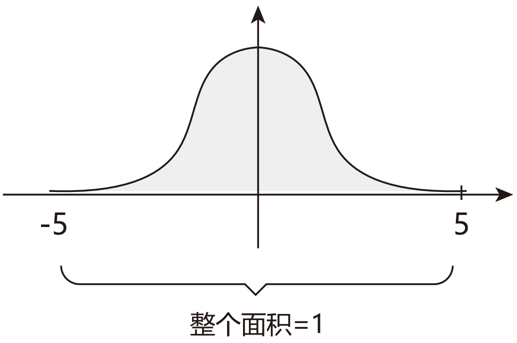
2.3.2 对于 $x\geq 5$ 的 y 值, 已经非常靠近 y=0 了

正态分布的值, 怎么算? — 查表.
一般, 书上给出的都是 "标准正态分布" 的表. 所以如果你是普通的 "正态分布", 必须先把它

转成“标准正态分布”, 再来查表.

并且, 表的范围, 只给出了 $0 \le x < 5$ 的值. **因为对于 $x \ge 5$ 的值, 此时的曲线高度, 即 y 值, 已经非常靠近 y=0 了. 所以我们就可以认为: 对于 $x \ge 5$ 的标准正态分布的“概率函数 $\varphi_0(x)$”的 y 值, 都 =0.**

同样, 来看累加函数 CDF:
对于 $x \ge 5$ 时, 其位置已经非常靠近整个曲线的右端末尾了, 而整个函数曲线下的面积也就 =1, 所以, 在 $x \ge 5$ 处的“累加函数 $\Phi_0(x)$”, 其值我们就可以认为是 1.



即:

标准正态分布	概率函数 $\varphi_0(x)$	累加函数 $\Phi_0(x)$
当 $x \le -5$ 时	$y \approx 0$	$y \approx 0$
当 $x \ge 5$ 时	$y \approx 0$	$y \approx 1$

3 普通的“正态分布”, 怎样转化成“标准正态分布”?

3.1 “概率函数”的转化公式是: $\boxed{\varphi(x)=\frac{1}{\sigma}\cdot\varphi_0(\frac{x-\mu}{\sigma})}$

普通“正态分布”的概率函数公式是:

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

我们来给它做一下变形(注意标出颜色的地方的变化):

$$\begin{aligned}\varphi(x)&=\frac{1}{\sqrt{2\pi}\cdot\sigma}e^{-\frac{(\frac{x-\mu}{\sigma})^2}{2}}\\&=\frac{1}{\sigma}\frac{1}{\sqrt{2\pi}}e^{-\frac{(\frac{x-\mu}{\sigma})^2}{2}}\end{aligned}$$

即: $\underbrace{\varphi(x)}_{\text{正态分布的概率函数}}=\frac{1}{\sigma}\frac{1}{\sqrt{2\pi}}e^{-\frac{(\frac{x-\mu}{\sigma})^2}{2}}$

变到这里后, 你来和“标准正态分布”的概率函数

$$\underbrace{\varphi_0(x)}_{\text{标准正态分布的概率函数}}=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$$

来做对比. 会发现: 两者的差别只在于红色标出的地方.

所以, 两者“概率函数”的转化公式, 就是:

$$\boxed{\underbrace{\varphi(x)}_{\text{正态分布的概率函数}}=\frac{1}{\sigma}\cdot\underbrace{\varphi_0(\overbrace{\frac{x-\mu}{\sigma}}^{\text{这一块, 就是标准正态分布的}x\text{值}})}_{\text{标准正态分布的概率函数}}}$$

3.2 “累加函数”的转化公式是: $\boxed{\Phi(x)=\Phi_0(\frac{x-\mu}{\sigma})}$

普通“正态分布”的“累加函数”公式是：

$$\Phi(x)=\frac{1}{\sqrt{2\pi}\cdot\sigma}\int_{-\infty}^{x}\left[e^{-\frac{(x-\mu)^2}{2\sigma^2}}\right]dx$$

我们来给它做一下变形（注意标出颜色的地方的变化）：

$$\begin{aligned}\Phi(x)&=\frac{1}{\sqrt{2\pi}\cdot\sigma}\int_{-\infty}^{x}\left[e^{-\frac{(\frac{x-\mu}{\sigma})^2}{2}}\right]dx\\&=\frac{1}{\sigma}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\left[e^{-\frac{(\frac{x-\mu}{\sigma})^2}{2}}\right]d(x-\underbrace{\mu}_{\text{常数}})\\&=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\left[e^{-\frac{(\frac{x-\mu}{\sigma})^2}{2}}\right]d\left(\frac{x-\mu}{\sigma}\right)\end{aligned}$$

即：$\underbrace{\Phi(x)}_{\text{正态分布的累加函数}}=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\left[e^{-\frac{(\frac{x-\mu}{\sigma})^2}{2}}\right]d\left(\frac{x-\mu}{\sigma}\right)$

变到这里后，你来和“标准正态分布”的“累加函数”

$$\underbrace{\Phi_0(x)}_{\text{标准正态分布的累加函数}}=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\left[e^{-\frac{x}{2}}\right]dx$$

来做对比．会发现：两者的差别只在于红色标出的地方．

所以，两者概率函数的转化公式，就是：

$$\boxed{\underbrace{\Phi(x)}_{\text{正态分布的累加函数}}=\underbrace{\Phi_0(\frac{x-\mu}{\sigma})}_{\text{标准正态分布的累加函数}}}$$

例

有 $X\sim N(1,4)$, 即 $\mu=1,\ \sigma^2=4,\ \sigma=2$

到注意: 这只是一个普通的正态分布, 我们必须先把它转成“标准正态分布”再来做.