第三章 一般随机变量

Probability and Statistics for Computer Scientists

第三章 一般随机变量

- 连续随机变量和概率密度函数
- 分布函数
- 正态随机变量
- 多个随机变量的联合概率密度
- 条件
- 连续贝叶斯准则

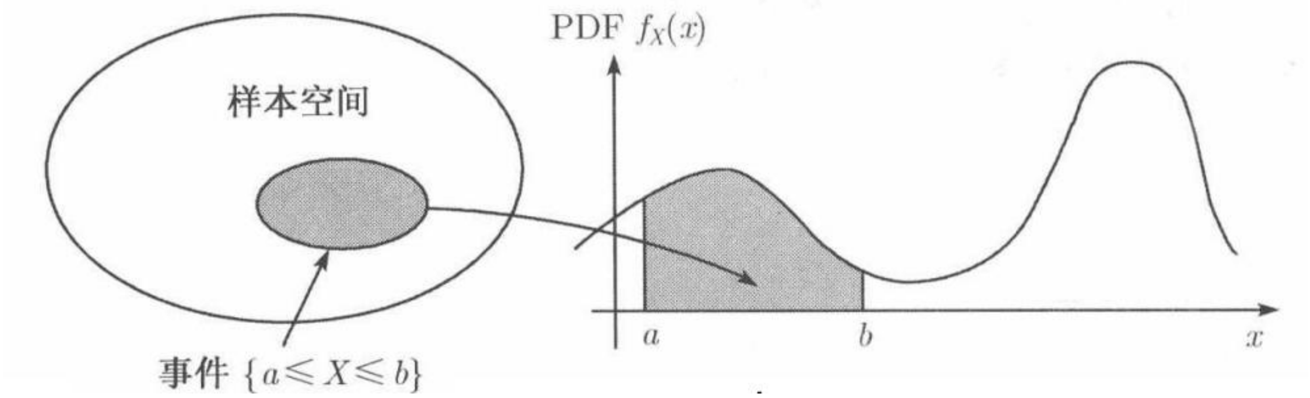
§1 连续随机变量和概率密度函数

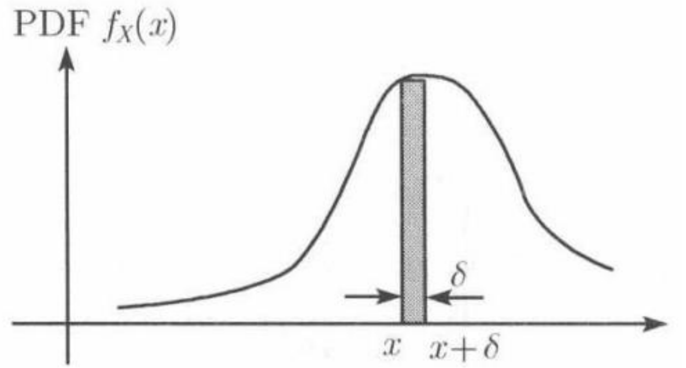
概率密度函数（PDF）：X落入x附近的单位长度的概率

$$\int_{\infty}^{\infty} f_X(x)\mathrm{d}x = \mathrm{P}(-\infty < X < \infty) = 1$$

$$\mathrm{P}(a \leqslant X \leqslant b) = \int_a^b f_X(x)\mathrm{d}x$$

$$\mathrm{P}([x, x+\delta]) = \int_x^{x+\delta} f_X(x)\mathrm{d}x \approx f_X(x) \cdot \delta$$





§1 连续随机变量和概率密度函数

例 3.3 (可以取任意大的值的 PDF) 考虑 X 的下列 PDF

$$f_X(x) = \begin{cases} \dfrac{1}{2\sqrt{x}}, & \text{若 } 0 < x \leqslant 1, \\ 0, & \text{其他}. \end{cases}$$

§1 连续随机变量和概率密度函数

关于 PDF 性质的小结

设 X 的 PDF(概率密度函数) 为 $f_X(x)$.

- $f_X(x) \geqslant 0$ 对一切 x 成立.
- $\int_{-\infty}^{\infty} f_X(x)\mathrm{d}x = 1$.
- 设 δ 是一个充分小的正数, 则 $\mathrm{P}([x, x+\delta]) \approx f_X(x) \cdot \delta$.
- 对任何实数轴上的子集 B,

$$\mathrm{P}(X \in B) = \int_B f_X(x)\mathrm{d}x.$$

§1 连续随机变量和概率密度函数

1.1 期望

$$\mathrm{E}[X]=\int_{-\infty}^{\infty} x f_X(x)\mathrm{d}x$$

- 关于随机变量 $g(X)$ 的期望规则为

$$\mathrm{E}[g(X)]=\int_{-\infty}^{\infty} g(x) f_X(x)\mathrm{d}x.$$

- X 的方差由下式给出：

$$\mathrm{var}(X)=\mathrm{E}[(X-\mathrm{E}[X])^2]=\int_{-\infty}^{\infty}(x-\mathrm{E}[X])^2 f_X(x)\mathrm{d}x.$$

- 关于方差，下列公式成立：

$$0\leqslant \mathrm{var}(X)=\mathrm{E}[X^2]-(\mathrm{E}[X])^2.$$

- 设$Y=aX+b$，其中 a 和 b 为常数，则

$$\mathrm{E}[Y]=a\mathrm{E}[X]+b,\quad \mathrm{var}(Y)=a^2\mathrm{var}(X).$$

§1 连续随机变量和概率密度函数

1.2 指数随机变量

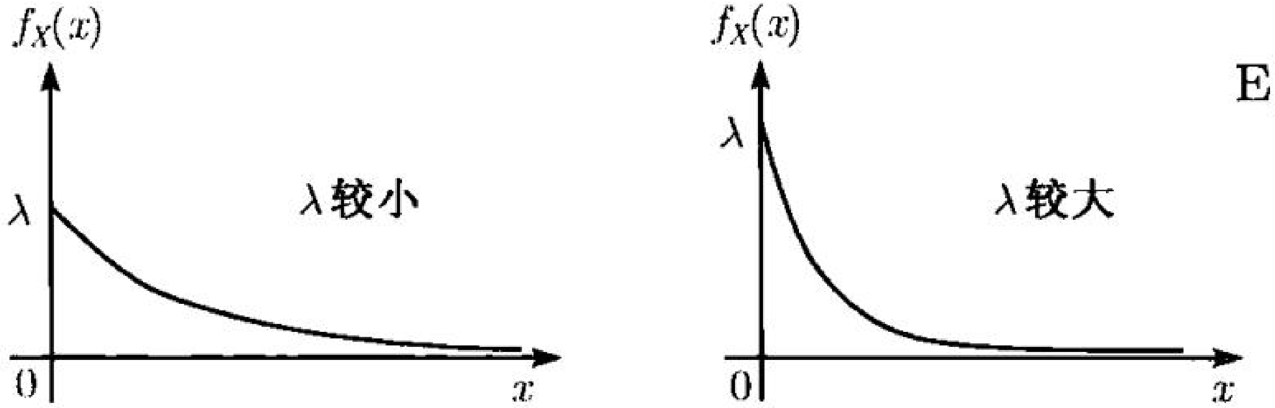
$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & 若\ x \geqslant 0, \\ 0, & 其他, \end{cases} \qquad E[X] = \frac{1}{\lambda}, \quad \mathrm{var}(X) = \frac{1}{\lambda^2}.$$

合法性

$$\int_{-\infty}^{\infty} f_X(x)\mathrm{d}x = \int_0^{\infty} \lambda e^{-\lambda x}\mathrm{d}x = -e^{-\lambda x}\Big|_0^{\infty} = 1.$$

§1 连续随机变量和概率密度函数

$$\mathrm{P}(X \geqslant a)=\int_a^{\infty} \lambda \mathrm{e}^{-\lambda x} \mathrm{d} x=-\left.\mathrm{e}^{-\lambda x}\right|_a^{\infty}=\mathrm{e}^{-\lambda a}.$$



$$\mathrm{E}[X]=\frac{1}{\lambda}, \quad \operatorname{var}(X)=\frac{1}{\lambda^2}.$$

§1 连续随机变量和概率密度函数

指数随机变量应用于科学观测

例 3.5 小陨石落入非洲撒哈拉沙漠的时间是遵从指数族分布的. 具体地说, 从某一观察者开始观察, 直到发现一颗陨石落到沙漠, 这个时间的分布是指数分布, 这个时间的平均长度是 10 天. 现在假定, 目前时间为晚上 12 点整. 问在第二天早晨 6:00 到傍晚 6:00 之间陨石首次落下的概率有多大?

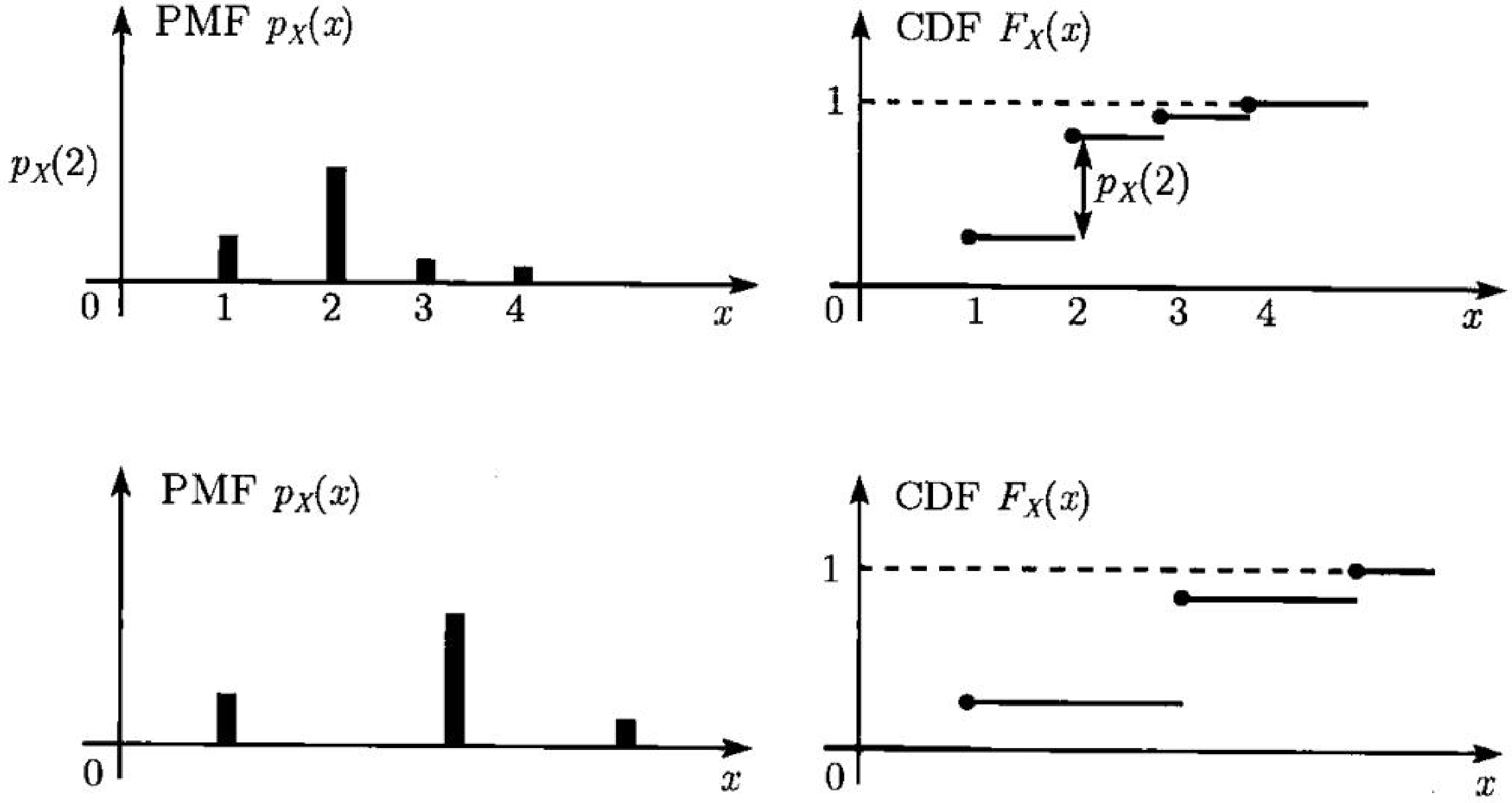
$$\mathrm{P}(1/4 \leqslant X \leqslant 3/4) = \mathrm{P}(X \geqslant 1/4) - \mathrm{P}(X > 3/4) = \mathrm{e}^{-\frac{1}{40}} - \mathrm{e}^{-\frac{3}{40}} = 0.047\,6,$$

§2 分布函数

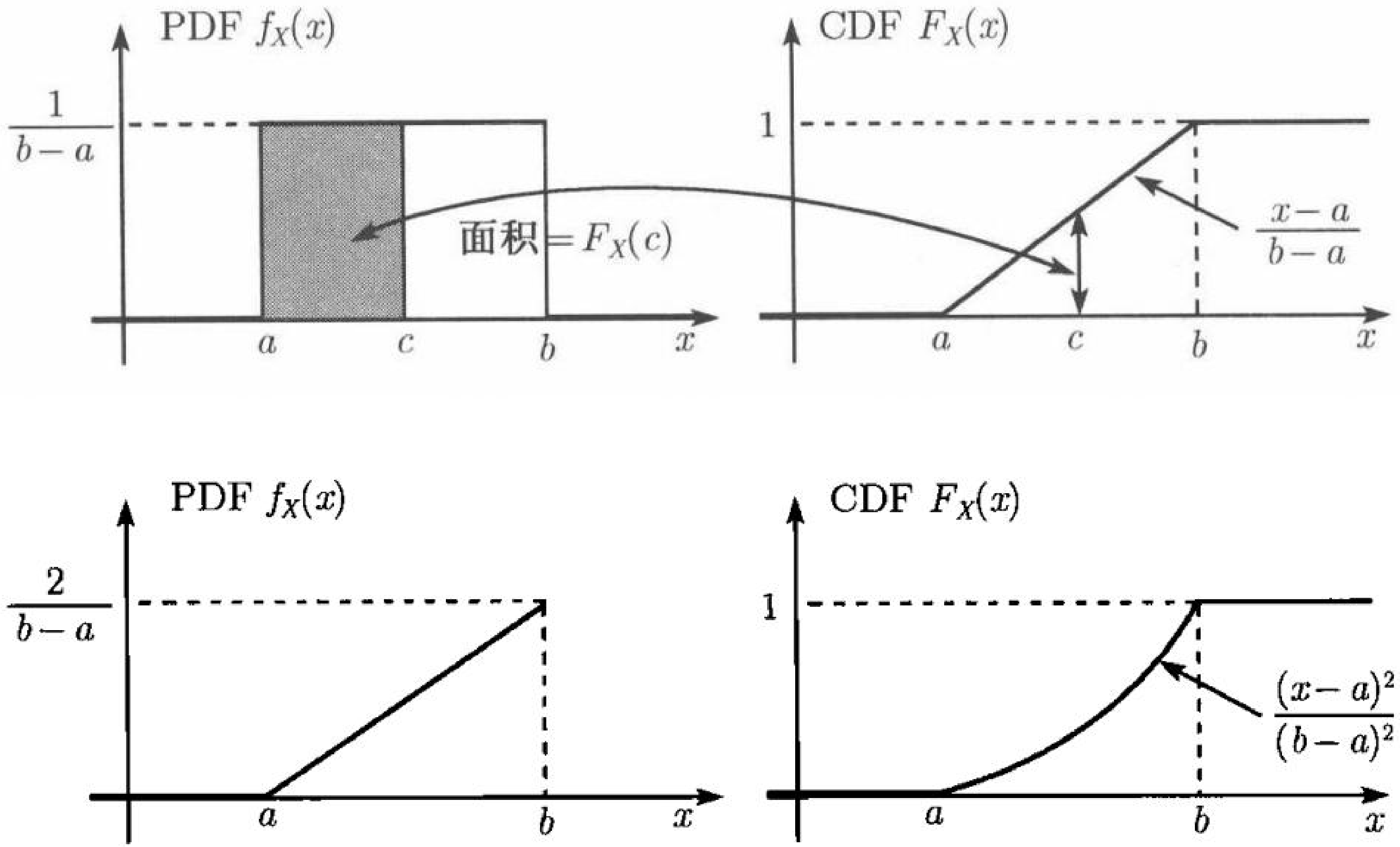
2.1 累积分布函数（CDF）：

$$F_X(x) = \mathrm{P}(X \leqslant x) = \begin{cases} \sum\limits_{k \leqslant x} p_X(k), & \text{若 } X \text{ 离散的}, \\ \int_{-\infty}^{x} f_X(t)\mathrm{d}t, & \text{若 } X \text{ 连续的}. \end{cases}$$

§2 分布函数



§2 分布函数



PDF $f_X(x)$
$\frac{1}{b-a}$
面积$=F_X(c)$
a c b x
CDF $F_X(x)$
1
$\frac{x-a}{b-a}$
a c b x
PDF $f_X(x)$
$\frac{2}{b-a}$
a b x
CDF $F_X(x)$
1
$\frac{(x-a)^2}{(b-a)^2}$
a b x

§ 2 分布函数

CDF 的性质

X 的 CDF $F_X(x)$ 是由下式定义的,

$$F_X(x) = \mathrm{P}(X \leqslant x), \quad \forall x,$$

并且 $F_X(x)$ 具有下列性质.

- $F_X(x)$ 是 x 的单调非减函数:

$$\text{若 } x \leqslant y, \text{则 } F_X(x) \leqslant F_X(y).$$

- 当 $x \to -\infty$ 的时候, $F_X(x)$ 趋于 0, 当 $x \to \infty$ 的时候, $F_X(x)$ 趋于 1.
- 当 X 是离散随机变量的时候, $F_X(x)$ 为阶梯函数.
- 当 X 是连续随机变量的时候, $F_X(x)$ 为 x 的连续函数.

§2 分布函数

- 当 X 是离散随机变量并且取整数值时, 分布函数和分布列可以利用求和或差分互求:

$$F_X(k) = \sum_{i=-\infty}^{k} p_X(i),$$

$$p_X(k) = \mathrm{P}(X \leqslant k) - \mathrm{P}(X \leqslant k-1) = F_X(k) - F_X(k-1),$$

其中 k 可以是任意整数.

- 当 X 是连续随机变量的时候, 分布函数和概率密度函数可以利用积分或微分互求:

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\mathrm{d}t, \quad f_X(x) = \frac{\mathrm{d}F_X}{\mathrm{d}x}(x).$$

(第二个等式只在分布函数可微的那些点上成立.)

§2 分布函数

2.2 考察几何和指数随机变量的分布函数

$$F_{\text{geo}}(n) = \sum_{k=1}^{n} p(1-p)^{k-1} = p\frac{1-(1-p)^n}{1-(1-p)} = 1-(1-p)^n, \quad n = 1, 2, \cdots.$$

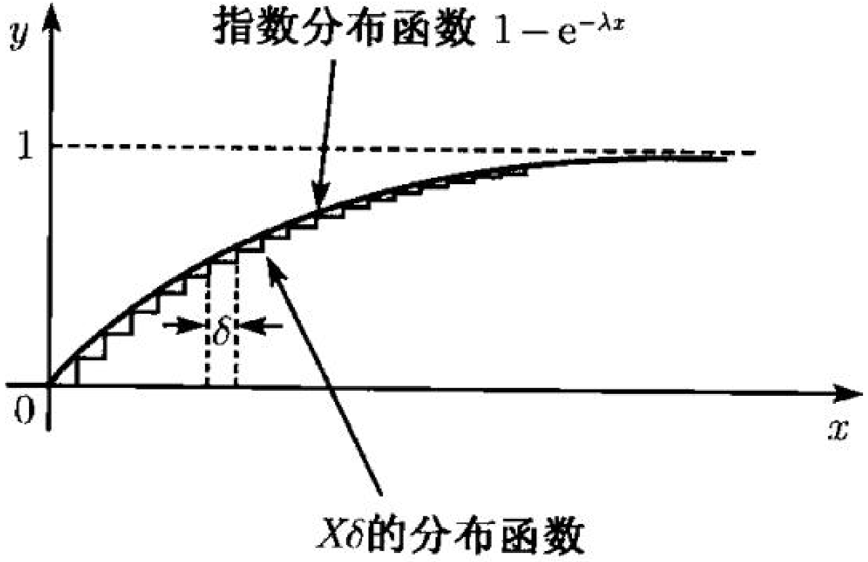
$$F_{\exp}(x) = \int_0^x \lambda \mathrm{e}^{-\lambda t} \mathrm{d}t = -\mathrm{e}^{-\lambda t}\Big|_0^x = 1 - \mathrm{e}^{-\lambda x}, \quad x > 0.$$

令 $\delta = -\ln(1-p)/\lambda$, $\mathrm{e}^{-\lambda\delta} = 1-p$.

分布函数 $F_{\exp}$ 在 $x = n\delta$ 处是与 F_{geo} 在 n 处相等的, $n = 1, 2, \cdots$, 即

$$F_{\exp}(n\delta) = F_{\text{geo}}(n), \quad n = 1, 2, \cdots.$$

§2 分布函数

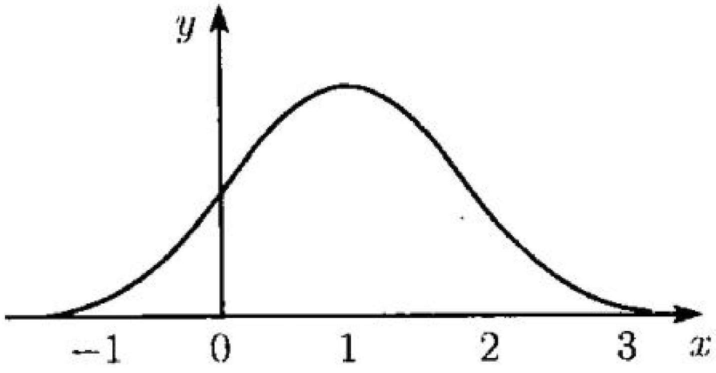


当 $\delta \to 0$ 时, $X\delta$ 的分布函数趋于指数分布函数 $1-\mathrm{e}^{\lambda x}$

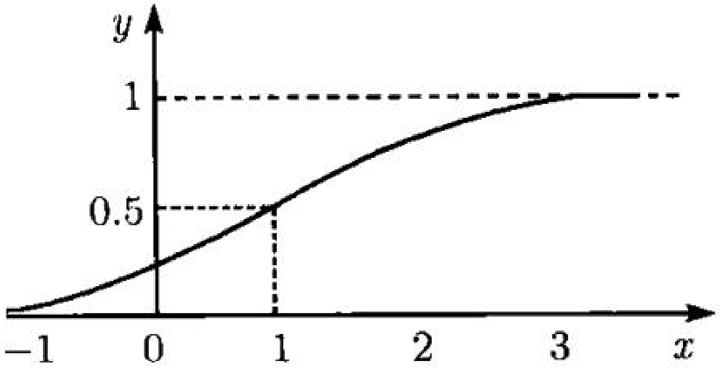
§3 正态随机变量

3.1 正态或高斯分布：

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



正态 PDF $f_X(x)$



正态 CDF $F_X(x)$

§3 正态随机变量

方差

$$\mathrm{var}(X) = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}(x-\mu)^2 \mathrm{e}^{-(x-\mu)^2/(2\sigma^2)}\mathrm{d}x.$$

积分变量替换 $y = (x-\mu)/\sigma$

$$\begin{aligned}
\mathrm{var}(X) &= \frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{\infty} y^2 \mathrm{e}^{-y^2/2}\mathrm{d}y \\
&= \frac{\sigma^2}{\sqrt{2\pi}}\left(-y\mathrm{e}^{-y^2/2}\right)\Big|_{-\infty}^{\infty} + \frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{e}^{-y^2/2}\mathrm{d}y \\
&= \frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{e}^{-y^2/2}\mathrm{d}y \\
&= \sigma^2.
\end{aligned}$$

§3 正态随机变量

线性变换之下随机变量的正态性保持不变

设 X 是正态随机变量, 其均值为 μ, 方差为 σ^2. 若 $a \neq 0$ 和 b 为两个常数, 则随机变量

$$Y = aX + b$$

仍然是正态随机变量, 其均值和方差由下式给出:

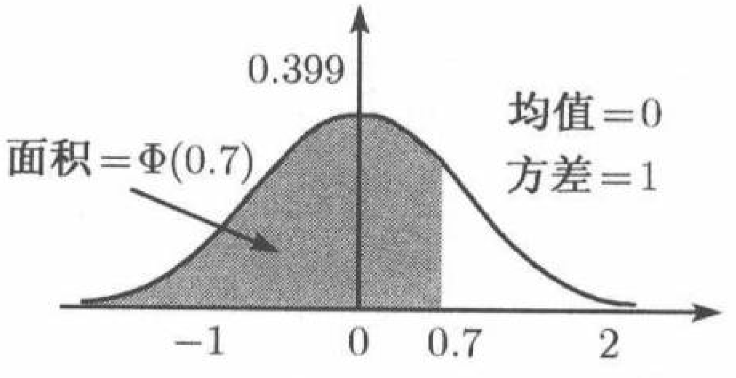
$$\mathrm{E}[Y] = a\mu + b, \quad \mathrm{var}(Y) = a^2\sigma^2.$$

§3 正态随机变量

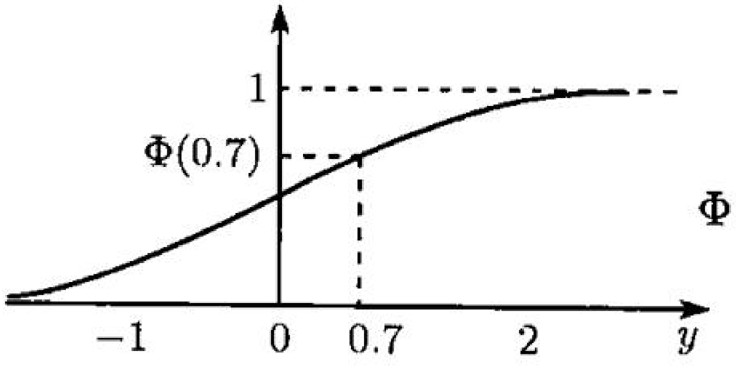
3.2 标准正态随机变量

$$\Phi(y)=P(Y\leqslant y)=P(Y<y)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y}e^{-t^2/2}dt.$$

$\Phi(-y)=1-\Phi(y)$, 对一切 y 成立.



标准正态 PDF



标准正态 CDF $\Phi(y)$

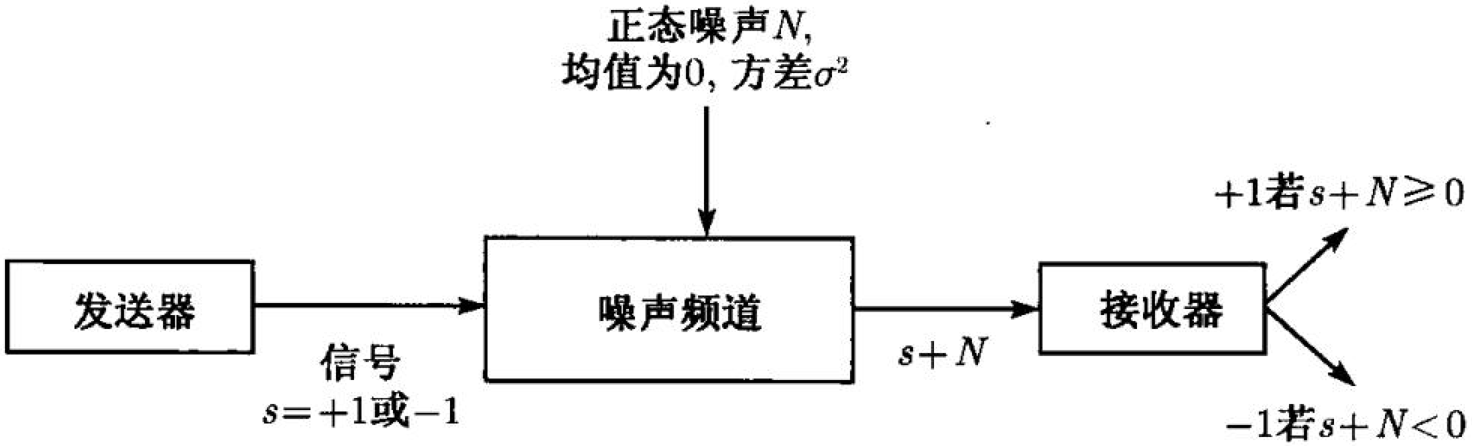
$\Phi(y)$ 的数值有表可查

§3 正态随机变量

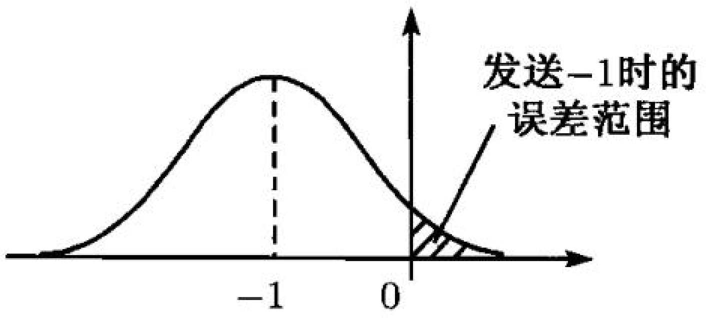
信号处理和通信工程中噪声的处理

例 3.8 (信号检测) 记一个传输的信号为 S, $S=1$ 或 $S=-1$. 由于通信误差, 在接收端得到的是加有噪声的信号, 噪声 N 是一个正态随机变量, 均值为 $\mu=0$, 方差为 σ^2. 如果接收端得到的混有噪声的信号大于 0, 则判断信号 $S=1$; 如果接收端得到的混有噪声的信号小于 0, 则判断信号 $S=-1$(见图 3.11). 问这种判断方法的误差有多大?

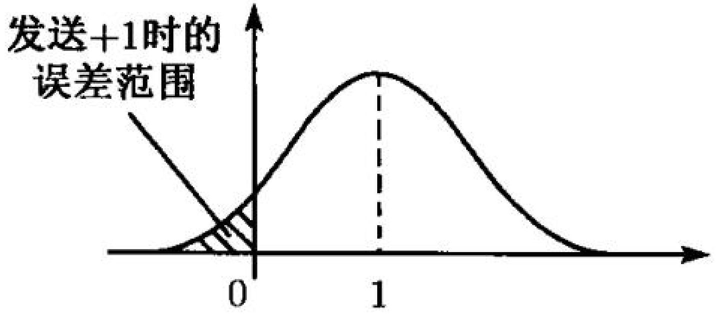
§3 正态随机变量



正态噪声N,
均值为0, 方差σ^2
发送器
信号
$s=+1$或-1
噪声频道
$s+N$
接收器
$+1$若$s+N\geqslant 0$
-1若$s+N<0$



发送−1时的
误差范围
−1
0



发送+1时的
误差范围
0
1

§3 正态随机变量

当传输方传输的信号为 $S = -1$, 而噪声 $N > 1$, 此时 $S + N = N - 1 > 0$,

当传输方传输的信号为 $S = 1$, 而噪声 $N < -1$, 此时 $S + N = N + 1 < 0$,

$$\begin{aligned} \mathrm{P}(N > 1) &= 1 - \mathrm{P}(N \leqslant 1) = 1 - \mathrm{P}\left(\frac{N-\mu}{\sigma} \leqslant \frac{1-\mu}{\sigma}\right) \\ &= 1 - \Phi\left(\frac{1-\mu}{\sigma}\right) = 1 - \Phi\left(\frac{1}{\sigma}\right). \end{aligned}$$

由对称性可知, 若发送的信号为 $S = +1$, 其相应的误判概率也是 $1 - \Phi(1/\sigma)$.

§4 多个随机变量的联合概率密度

4.1 连续随机变量的联合概率密度函数

$$\mathrm{P}((X,Y)\in B)=\int\int_{(x,y)\in B} f_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y.$$

$$F_{X,Y}(x,y)=\mathrm{P}(X\leqslant x, Y\leqslant y)=\int_{-\infty}^{x}\int_{-\infty}^{y} f_{X,Y}(s,t)\mathrm{d}s\mathrm{d}t.$$

密度函数的归一化条件

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y=1.$$

联合概率密度函数的意义

$$\mathrm{P}(a\leqslant X\leqslant a+\delta, c\leqslant Y\leqslant c+\delta)=\int_{a}^{a+\delta}\int_{c}^{c+\delta} f_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y\approx f_{X,Y}(a,c)\cdot\delta^2$$

§4 多个随机变量的联合概率密度

4.2 连续随机变量的边缘概率密度

计算单独一个随机变量 (X 或 Y) 所刻画的事件的概率

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\mathrm{d}y.$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\mathrm{d}x.$$

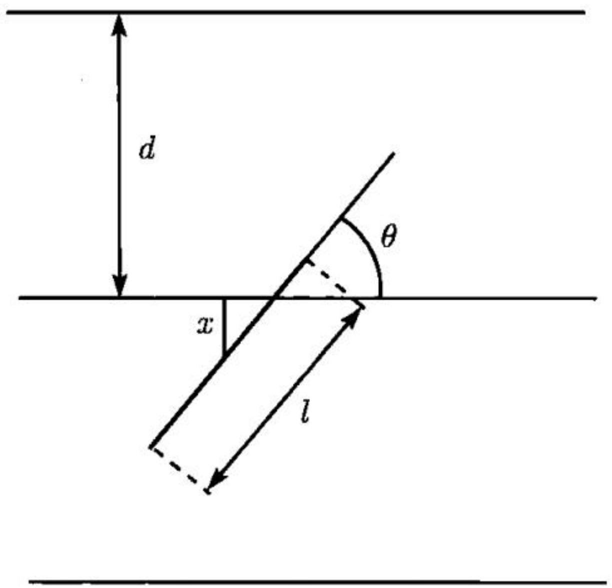
考虑

$$\mathrm{P}(X \in A) = \mathrm{P}(X \in A, Y \in (-\infty,\infty)) = \int_A \int_{-\infty}^{\infty} f_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y = \int_A f_X(x)\mathrm{d}x$$

§4 多个随机变量的联合概率密度

例 3.11 (蒲丰的抛针试验)[①] 这是一个著名的例子, 几何概率由此发源. 所讨论的问题是对随机放置的对象的几何性质的分析.

在平面上画了若干条平行线, 相互之间的距离为 d 现在往平面上随机地抛掷一根针, 针的长度为 l. 问针与直线相交的概率有多大?



§4 多个随机变量的联合概率密度

我们假定 $l < d$, 这样针不能同时与两条直线同时相交. 令 X 为针的中点离最近的那一条直线的垂直距离, Θ 表示针与平行直线之间的夹角 (见图 3.13). 我们假定 (X,Θ) 的联合概率密度函数为矩形集合 $\{(x,\theta)|0 \leqslant x \leqslant d/2, 0 \leqslant \theta \leqslant \pi/2\}$ 上的联合均匀概率密度函数.

$$f_{X,\Theta}(x,\theta) = \begin{cases} 4/(\pi d), & \text{若 } x \in [0,d/2] \text{ 和 } \theta \in [0,\pi/2], \\ 0, & \text{其他}. \end{cases}$$

针与平行直线相交的充要条件为 $X \leqslant \dfrac{l}{2}\sin\Theta$,

$$\mathrm{P}(X \leqslant (l/2)\sin\Theta) = \int\int_{x \leqslant (l/2)\sin\theta} f_{X,\Theta}(x,\theta)\,\mathrm{d}x\mathrm{d}\theta = \frac{2l}{\pi d}.$$

§4 多个随机变量的联合概率密度

4.3 连续随机变量的联合分布函数

$$F_{X,Y}(x,y) = \mathrm{P}(X \leqslant x, Y \leqslant y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{X,Y}(s,t)\mathrm{d}s\mathrm{d}t.$$

4.4 多个连续随机变量的期望

$$\mathrm{E}[g(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) f_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y.$$

$$\mathrm{E}[aX + bY + c] = a\mathrm{E}[X] + b\mathrm{E}[Y] + c$$

§4 多个随机变量的联合概率密度

4.5 多于两个连续随机变量的情况

$$\mathrm{P}((X,Y,Z)\in B)=\iiint_{(x,y,z)\in B} f_{X,Y,Z}(x,y,z)\,\mathrm{d}x\mathrm{d}y\mathrm{d}z.$$

联合概率密度函数

$$f_{X,Y}(x,y)=\int_{-\infty}^{\infty} f_{X,Y,Z}(x,y,z)\,\mathrm{d}z,$$

$$f_X(x)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y,Z}(x,y,z)\,\mathrm{d}y\mathrm{d}z.$$

随机变量 $g(X,Y,Z)$ 的期望

$$\mathrm{E}[g(X,Y,Z)]=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y,z)f_{X,Y,Z}(x,y,z)\,\mathrm{d}x\mathrm{d}y\mathrm{d}z.$$

§5 条件

5.1 连续情况下以事件为条件的随机变量

条件概率密度函数：$\mathrm{P}(X \in B|A) = \int_B f_{X|A}(x)\mathrm{d}x.$

归一化等式 $\int_{-\infty}^{\infty} f_{X|A}(x)\mathrm{d}x = 1,$

对比：$\mathrm{P}(X \in B|X \in A) = \dfrac{\mathrm{P}(X \in A, X \in B)}{\mathrm{P}(X \in A)} = \dfrac{\int_{A\cap B} f_X(x)\mathrm{d}x}{\mathrm{P}(X \in A)}$

可知 $f_{X|A}(x) = \begin{cases} \dfrac{f_X(x)}{\mathrm{P}(X \in A)}, & \text{若 } x \in A, \\ 0, & \text{其他.} \end{cases}$

§5 条件

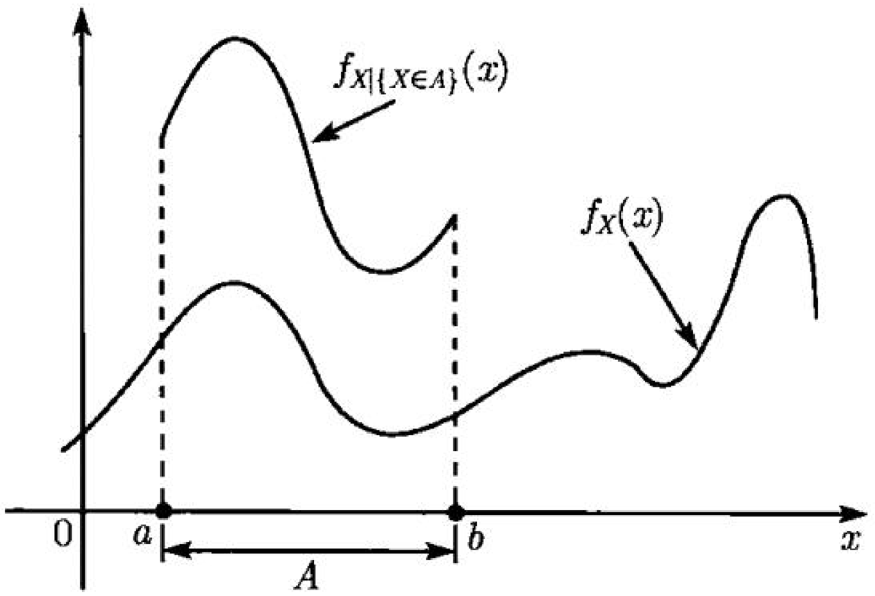
运用条件概率：讨论指数函数的无记忆性

例 3.13 (指数随机变量的无记忆性) 一个灯泡的使用寿命 T 是一个指数随机变量, 其参数为 λ. 阿丽将灯打开后离开房间, 在外面呆了一段时间以后 (时间长度为 t), 她回到房间, 灯还是亮着. 这相当于事件 $A=\{T>t\}$ 发生了. 记 X 为灯泡的剩余寿命, 问 X 的分布函数是什么?

§5 条件

$$
\begin{aligned}
P(X > x|A) &= P(T > t + x|T > t) \\
&= \frac{P(T > t + x \text{ 且 } T > t)}{P(T > t)} \\
&= \frac{P(T > t + x)}{P(T > t)} \\
&= \frac{e^{-\lambda(t+x)}}{e^{-\lambda t}} \\
&= e^{-\lambda x}
\end{aligned}
$$

§5 条件



若完成某个任务的时间服从指数分布，只要这个任务没完成，剩余时间仍服从指数分布且参数不变

§5 条件

5.2 条件概率密度下的全概率定理

设 $A_1, A_2, \cdots, A_n$ 为互不相容的 n 个事件，对每个 i, $\mathrm{P}(A_i) > 0$, 并且这些事件形成样本空间的一个分割. 则

$$f_X(x) = \sum_{i=1}^{n} \mathrm{P}(A_i) f_{X|A_i}(x)$$

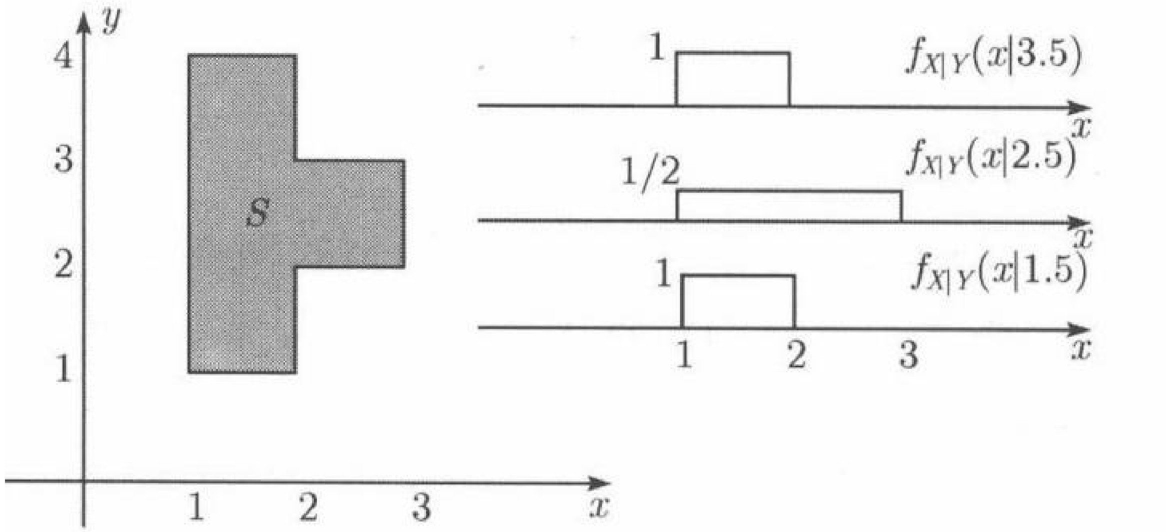
(全概率公式的一种变形).

§5 条件

5.3 连续情况下一个随机变量对另一个随机变量的条件

在给定 $Y=y$ 的情况下, X 的条件概率密度函数

$$f_{X|Y}(x|y)=\frac{f_{X,Y}(x,y)}{f_Y(y)}.$$



§5 条件

5.4 连续随机变量的条件期望

条件期望性质的小结

记 X 和 Y 为联合连续的随机变量，A 是满足 $\mathrm{P}(A)>0$ 的事件.

- X 在给定事件 A 之下的条件期望由下式定义

$$\mathrm{E}[X|A]=\int_{-\infty}^{\infty} x f_{X|A}(x)\mathrm{d}x,$$

给定 $Y=y$ 之下的条件期望由下式定义

$$\mathrm{E}[X|Y=y]=\int_{-\infty}^{\infty} x f_{X|Y}(x|y)\mathrm{d}x.$$

- 期望规则仍然有效：

$$\mathrm{E}[g(X)|A]=\int_{-\infty}^{\infty} g(x) f_{X|A}(x)\mathrm{d}x,$$

$$\mathrm{E}[g(X)|Y=y]=\int_{-\infty}^{\infty} g(x) f_{X|Y}(x|y)\mathrm{d}x.$$

§ 5 条件

- **全期望定理**: 设 $A_1, A_2, \cdots, A_n$ 为互不相容的 n 个事件, 对每个 i, $\mathrm{P}(A_i) > 0$, 并且这些事件形成样本空间的一个分割. 则

$$\mathrm{E}[X] = \sum_{i=1}^{n} \mathrm{P}(A_i)\mathrm{E}[X|A_i].$$

相似地,

$$\mathrm{E}[X] = \int_{-\infty}^{\infty} \mathrm{E}[X|Y=y] f_Y(y) \mathrm{d}y.$$

- 涉及几个随机变量的函数的情况, 具有完全相似的结果. 例如

$$\mathrm{E}[g(X,Y)|Y=y] = \int g(X,Y) f_{X|Y}(x|y) \mathrm{d}x,$$

$$\mathrm{E}[g(X,Y)] = \int \mathrm{E}[g(X,Y)|Y=y] f_Y(y) \mathrm{d}y.$$

§ 5 条件

全期望定理的验证：

展开.

$$
\begin{aligned}
\int_{-\infty}^{+\infty} E[X|Y=y]\, f_Y(y)\,dy &= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} x f_{X|Y}(x|y)\,dx\right] f_Y(y)\,dy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\, f_{X|Y}(x|y)\, f_Y(y)\,dx\,dy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\, f_{X,Y}(x,y)\,dx\,dy \\
&= \int_{-\infty}^{\infty} x\left[\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy\right] dx \\
&= \int_{-\infty}^{\infty} x\, f_X(x)\,dx \\
&= E[X]
\end{aligned}
$$

§ 5 条件

5.5 连续随机变量的独立性

$$f_{X,Y}(x,y) = f_X(x)f_Y(y), \quad \text{对一切 } x, y \text{ 成立.}$$

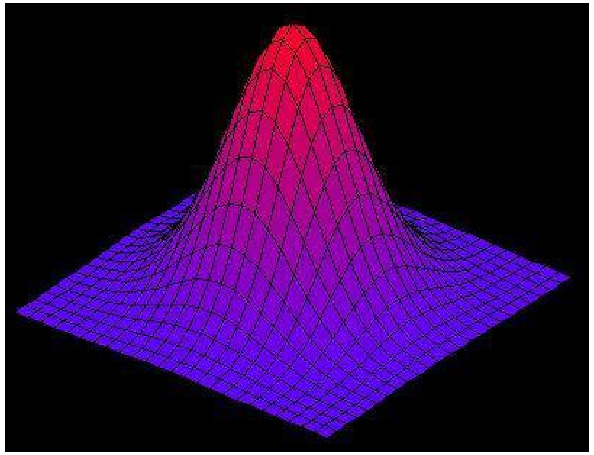
$$f_{X,Y,Z}(x,y,z) = f_X(x)f_Y(y)f_Z(z), \quad \text{对一切} x, y, z \text{成立.}$$

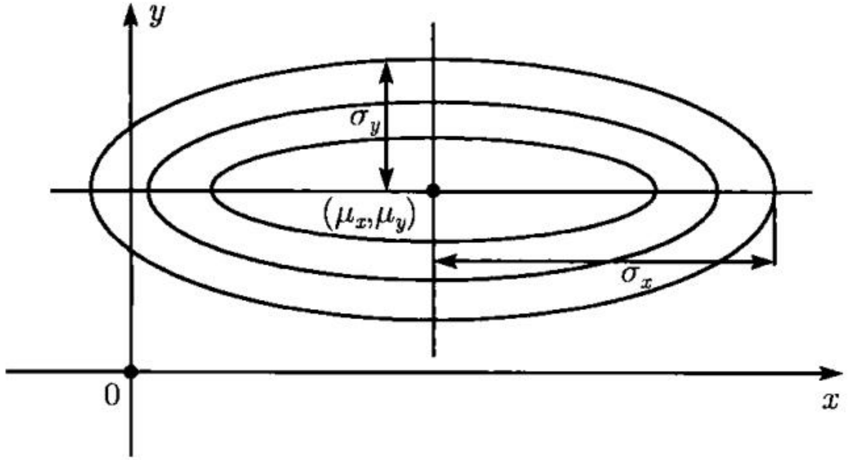
讨论独立的正态随机变量

例 3.18 (独立的正态随机变量) 设 X 和 Y 是相互独立的正态随机变量, 其期望和方差分别为 μ_x, μ_y 和 σ_x^2, σ_y^2. 它们的联合分布密度函数为

$$f_{X,Y}(x,y) = f_X(x)f_Y(y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left\{-\frac{(x-\mu_x)^2}{2\sigma_x^2} - \frac{(y-\mu_y)^2}{2\sigma_y^2}\right\}.$$

§5 条件





$$\frac{(x-\mu_x)^2}{2\sigma_x^2}+\frac{(y-\mu_y)^2}{2\sigma_y^2}=\text{常数}.$$

§5 条件

若 X 和 Y 相互独立,

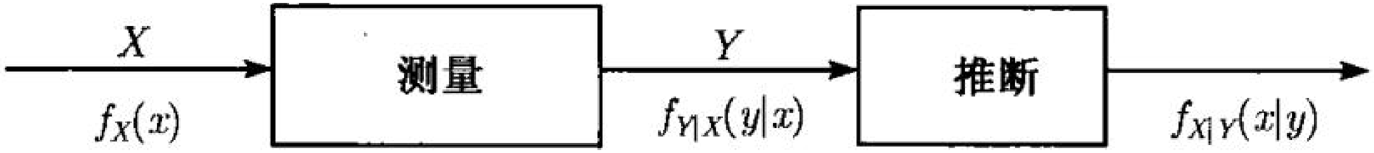
$$\begin{aligned}
\mathrm{P}(X \in A \text{ 和 } Y \in B) &= \int_{x\in A}\int_{y\in B} f_{X,Y}(x,y)\mathrm{d}y\mathrm{d}x \\
&= \int_{x\in A}\int_{y\in B} f_X(x)f_Y(y)\mathrm{d}y\mathrm{d}x \\
&= \int_{x\in A} f_X(x)\mathrm{d}x \int_{y\in B} f_Y(y)\mathrm{d}y \\
&= \mathrm{P}(X \in A)\mathrm{P}(Y \in B).
\end{aligned}$$

特别地, 独立性蕴涵

$$F_{X,Y}(x,y) = \mathrm{P}(X \leqslant x, Y \leqslant y) = \mathrm{P}(X \leqslant x)\mathrm{P}(Y \leqslant y) = F_X(x)F_Y(y).$$

§6 连续贝叶斯准则

6.1 连续随机变量的推断问题



$$f_{X|Y}(x|y) = \frac{f_X(x)f_{Y|X}(y|x)}{f_Y(y)} = \frac{f_X(x)f_{Y|X}(y|x)}{\int_{-\infty}^{\infty} f_X(t)f_{Y|X}(y|t)\mathrm{d}t}$$

§6 连续贝叶斯准则

由于事件 $\{Y=y\}$ 是一个零概率事件，我们转而考虑事件 $\{y \leqslant Y \leqslant y+\delta\}$ 然后令 δ 趋向于 0

$$
\begin{aligned}
\mathrm{P}(A|Y=y) &\approx \mathrm{P}(A|y \leqslant Y \leqslant y+\delta) \\
&= \frac{\mathrm{P}(A)\mathrm{P}(y \leqslant Y \leqslant y+\delta|A)}{\mathrm{P}(y \leqslant Y \leqslant y+\delta)} \\
&\approx \frac{\mathrm{P}(A) f_{Y|A}(y)\delta}{f_Y(y)\delta} \\
&= \frac{\mathrm{P}(A) f_{Y|A}(y)}{f_Y(y)}. \\
&= \frac{\mathrm{P}(A) f_{Y|A}(y)}{\mathrm{P}(A) f_{Y|A}(y) + \mathrm{P}(A^c) f_{Y|A^c}(y)}
\end{aligned}
$$

§6 连续贝叶斯准则

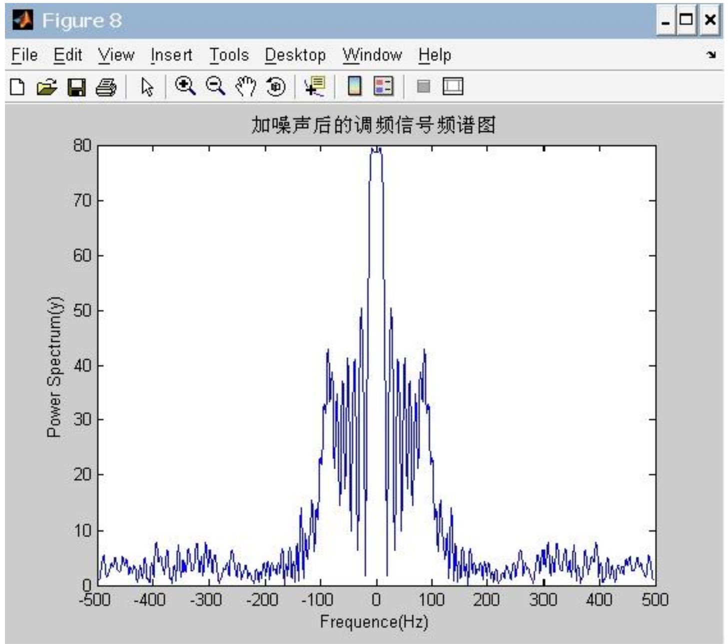
令事件 A 具有形式 $\{N=n\}$

$$\mathrm{P}(N=n|Y=y)=\frac{p_N(n)f_{Y|n}(y|n)}{f_Y(y)}$$

$$=\frac{p_N(n)f_{Y|n}(y|n)}{\sum_i p_N(i)f_{Y|N}(y|i)}$$

§6 连续贝叶斯准则

6.2 离散随机变量的推断



$$P(A|Y=y)=\frac{P(A)f_{Y|A}(y)}{P(A)f_{Y|A}(y)+P(A^c)f_{Y|A^c}(y)}.$$

$$P(N=n|Y=y)=\frac{p_N(n)f_{Y|n}(y|n)}{\sum_i p_N(i)f_{Y|N}(y|i)}.$$

§6 连续贝叶斯准则

讨论：信号检测的应用

例 3.20 (信号检测) 设 S 是一个只取两个值的信号. 记 $\mathrm{P}(S=1)=p$ 和 $\mathrm{P}(S=-1)=1-p$. 在接收端, 得到的信号为 $Y=N+S$, 其中 N 是一个正态噪声, 期望为 0, 方差为 1, 并且与 S 相互独立. 当观察到的信号为 y 的时候, $S=1$ 的概率是多少?

§6 连续贝叶斯准则

对于给定的 $S=s$, Y 是一个正态随机变量, 期望为 s, 方差为 1. 应用刚才得到的公式

$$\mathrm{P}(S=1|Y=y)=\frac{p_s(1)f_{Y|S}(y|1)}{f_Y(y)}=\frac{\frac{p}{\sqrt{2\pi}}\mathrm{e}^{-(y-1)^2/2}}{\frac{p}{\sqrt{2\pi}}\mathrm{e}^{-(y-1)^2/2}+\frac{1-p}{\sqrt{2\pi}}\mathrm{e}^{-(y+1)^2/2}},$$

将上式简化得

$$\mathrm{P}(S=1|Y=y)=\frac{p\mathrm{e}^{y}}{p\mathrm{e}^{y}+(1-p)\mathrm{e}^{-y}}.$$

§6 连续贝叶斯准则

6.3 基于离散观察值的推断

现在观察值是离散的.

$$f_{Y|A}(y) = \frac{f_A(y)\mathrm{P}(A|Y=y)}{\mathrm{P}(A)}.$$

$$f_{Y|A}(y) = \frac{f_A(y)\mathrm{P}(A|Y=y)}{\int_{-\infty}^{\infty} f_Y(t)\mathrm{P}(A|Y=t)\mathrm{d}t}.$$

本章介绍的几个概率模型总结

连续随机变量的某些结果

$[a,b]$上的连续均匀随机变量

$$f_X(x)=\begin{cases}\dfrac{1}{b-a}, & 若\ a\leqslant x\leqslant b,\\ 0, & 其他,\end{cases}$$

$$\mathrm{E}[X]=\frac{a+b}{2},\quad \mathrm{var}(X)=\frac{(b-a)^2}{12}.$$

分布参数为 λ 的指数随机变量

$$f_X(x)=\begin{cases}\lambda e^{-\lambda x}, & 若\ x\geqslant 0,\\ 0, & 其他,\end{cases}\qquad F_X(x)=\begin{cases}1-e^{-\lambda x}, & 若\ x\geqslant 0,\\ 0, & 其他,\end{cases}$$

$$\mathrm{E}[X]=\frac{1}{\lambda},\quad \mathrm{var}(X)=\frac{1}{\lambda^2}.$$

分布参数为 μ 和 σ^2 的正态随机变量

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

$$\mathrm{E}[X]=\mu,\quad \mathrm{var}(X)=\sigma^2.$$