

动机

• 单位时间内, 事件“平均”发生 λ 次.

X = 单位时间内 发生次数

泊松分布

$$P(X=k) = \frac{1}{k!} \lambda^k \cdot e^{-\lambda}$$

Y = 两次事件发生的间隔

$$\begin{aligned} P(Y \geq 1) &= P(X=0) = \frac{1}{0!} \cdot \lambda^0 \cdot e^{-\lambda} \\ &= e^{-\lambda} \end{aligned}$$

$$P(Y \geq t) = P(X_t=0) = e^{-t\lambda} \quad (t \geq 0)$$

原单位 $1d \rightsquigarrow$ 新单位 td

事件平均发生次数 $t\lambda$

X_t : td \sim 泊松分布 $\frac{1}{k!} (t\lambda)^k \cdot e^{-t\lambda}$

$$P(X < t) = 1 - P(X \geq t) = 1 - e^{-t\lambda}$$

$$\text{分布密度 } p(t) = \begin{cases} \frac{d}{dt} P(X < t) = \lambda \cdot e^{-\lambda t} & t \geq 0 \\ 0 & t < 0 \end{cases}$$

指数分布

若连续型随机变量 X 的分布密度函数为

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{当 } x > 0, \\ 0, & \text{当 } x \leq 0 \end{cases}$$

其中 $\lambda > 0$ 为参数, 记为 $X \sim \exp(\lambda)$, 其分布函数为

$$F_X(x) = \begin{cases} 1 - e^{-\lambda x}, & \text{当 } x > 0 \\ 0, & \text{当 } x \leq 0 \end{cases}$$

例子

某窗口接待一位顾客的服务时间 T 服从参数为 $\frac{1}{10}$ 的指数分布, 即

$$f_T(t) = \begin{cases} \frac{1}{10} e^{-\frac{1}{10}t}, & \text{当 } t > 0, \\ 0, & \text{当 } t \leq 0 \end{cases}$$

例子

假设一次服务时间超过 15 分钟, 顾客即评价为“不满意”, 求

- ▶ 某位顾客不满意的概率
- ▶ 十位顾客中恰有两位顾客评价为不满意的概率
- ▶ 十位顾客中最多有两位评价为不满意的概率
- ▶ 十位顾客中至少有两位评价为不满意的概率

例子

设事件 $A = \{\text{某位顾客不满意}\}$

$$P(A) = \int_{15}^{\infty} \frac{1}{10} e^{-\frac{1}{10}x} dx$$

由于每位顾客接受服务的试验相互独立, 所以某一窗口服务多位顾客是 n 重伯努利试验
记不满意的顾客数为 Y , $Y \sim B(10, P(A))$
那么题目中的第 2, 3, 4 小问分别对应下面三个概率

$$P(Y = 2), P(Y \leq 2), P(Y \geq 2)$$

例子

根据二项分布公式

$$P(Y=2) = \binom{10}{2} P(A)^2 P(\bar{A})^8$$

$$P(Y \leq 2) = \sum_{k=0}^2 \binom{10}{k} P(A)^k P(\bar{A})^{10-k}$$

$$\begin{aligned} P(Y \geq 2) &= 1 - P(Y \leq 1) \\ &= 1 - \sum_{k=0}^1 \binom{10}{k} P(A)^k P(\bar{A})^{10-k} \end{aligned}$$

指数分布的无记忆性

指数分布也具有无记忆性, 即

$$P(X > t + s | X > s) = P(X > t)$$

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

正态分布

正态分布

连续型随机变量服从的最常见的分布是 **正态分布**

正态分布的密度函数

若 X 服从正态分布, 那么其分布密度函数为

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

其中 $\mu \in \mathbb{R}$ 和 $\sigma > 0$ 为正态分布的参数, 记

$X \sim N(\mu, \sigma^2)$. *normal*

特别的, 当 $\mu = 0, \sigma = 1$ 时, 称 X 服从 **标准正态分布**.

标准正态分布

当 $X \sim N(0, 1)$, 其密度函数为

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

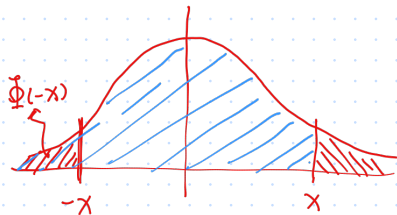
其分布函数为 $\Phi(x)$,

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

标准正态分布

显然

$$\begin{aligned}\Phi(-x) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-x} e^{-\frac{t^2}{2}} dt = -\frac{1}{\sqrt{2\pi}} \int_{\infty}^x e^{-\frac{\xi^2}{2}} d\xi \\ &= \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-\frac{\xi^2}{2}} d\xi = 1 - \Phi(x)\end{aligned}$$



标准正态分布

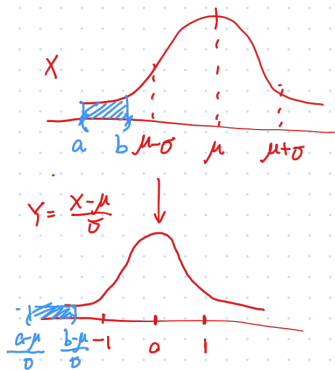
$$\begin{aligned}P(a < X < b) &= \frac{1}{\sqrt{2\pi}\sigma} \int_a^b e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\&= \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{(x-\mu)^2}{2\sigma^2}} d\left(\frac{x-\mu}{\sigma}\right) \\&= \frac{1}{\sqrt{2\pi}} \int_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}} e^{-\frac{t^2}{2}} dt \\&= \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)\end{aligned}$$

上面的 $\Phi\left(\frac{b-\mu}{\sigma}\right)$ 和 $\Phi\left(\frac{a-\mu}{\sigma}\right)$ 都可以查表计算.
对于 $X \sim N(\mu, \sigma^2)$, 查表计算有

$$P(X - \mu \leq 3\sigma) = 0.9973$$

命题 2.3.1

设 $X \sim N(\mu, \sigma^2)$, 令 $Y = \frac{X-\mu}{\sigma}$, 那么 $Y \sim N(0, 1)$.
证明稍后给出.



例 2.3.3

$$X \sim N(\mu, \sigma^2) \quad Y = \frac{X - \mu}{\sigma} = \frac{X - (-1)}{2} \sim N(0, 1^2)$$

设 $X \sim N(-1, 2^2)$, 求 $P(-5 \leq X < 1)$,
 $P(-2 \leq X \leq 2)$, $P(X \geq 1)$.

解:

$$\begin{aligned} P(-5 \leq X < 1) &= \Phi\left(\frac{1 - (-1)}{2}\right) - \Phi\left(\frac{-5 - (-1)}{2}\right) \\ &= \Phi(1) - \Phi(-2) \end{aligned}$$

$$\begin{aligned} P(-2 \leq X \leq 2) &= \Phi\left(\frac{2 - (-1)}{2}\right) - \Phi\left(\frac{-2 - (-1)}{2}\right) \\ &= \Phi\left(\frac{3}{2}\right) - \Phi\left(-\frac{1}{2}\right) \end{aligned}$$

$$P(X \geq 1) = 1 - P(X < 1)$$

例 2.3.4

假设一次测量误差 $X \sim N(0, 10^2)$, 现在独立重复的进行 100 次测量, 求误差的绝对值超过 19.6 的次数不小于 3 的概率.

解: 先求出一次测量误差大于 19.6 的概率.

$$P(|X| > 19.6) = 1 - P(|X| \leq 19.6)$$


$$= 1 - P\left(\left|\frac{X-0}{10}\right| \leq 1.96\right)$$

$$= 1 - [\Phi(1.96) - \Phi(-1.96)]$$

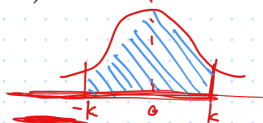
$$= 2 - 2\Phi(1.96) \approx 0.05$$

$$\begin{aligned} Y &= \frac{X-\mu}{\sigma} \\ &= \frac{X-0}{10} \end{aligned}$$

X


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

$$P(|X| \leq k) = P(X \leq k) - P(X < -k)$$



例 2.3.4

由于每次测量都是独立的, 所以设 Y 为 100 次测量中误差大于 19.6 的次数, 那么

$$Y \sim B(100, 0.05),$$

$$P(Y \geq 3) = 1 - P(Y \leq 2) = 1 - \sum_{k=0}^2 b(k; 100, 0.05) = 0.88$$

大家也可用泊松分布计算下看看.

例 2.3.5

公共汽车车门的高度是按照男子与车门顶碰头的机会 0.01 以下来设计的. 设男子的身高 $X \sim N(170, 6^2)$. 确定车门高度.

例 2.3.5

解：设车门高度为 h , 按照题意,

$$P(X > h) < 0.01,$$

所以 γ_2

$$\Phi\left(\frac{h - 170}{6}\right) = P(X \leq h) = 1 - P(X > h) > 0.99$$

查一下正态分布表, 近似的

$$\frac{h - 170}{6} \geq 2.33$$

即

$$h \geq 183.98$$

如果取整数的话, $h = 184$ cm.

例 2.3.6

设某只股票的初始价格为 $S_0 = 40$ 元, 预期收益率 μ 为每年 16%, 波动率为每年 20%. 在 Black-Scholes 模型下, 股票在每个时刻 t 的价格 S_t 为随机变量

$$S_t = S_0 \exp \left(\left(\mu - \frac{\sigma^2}{2} \right) t + \sigma B_t \right)$$

其中 $B_t \sim N(0, t)$, 估计六个月后这只股票的价格范围 (允许出错的概率为 5%) .

例 2.3.6



例 2.3.6

解：对于每个给定的时间 t , 随机变量 S_t 为另一个随机变量 B_t 的函数, 易知的是在 $t = 0.5$ (半年) 的时候

$$S_t = S_0 \cdot \exp(\mu \cdot t + \sigma \cdot B_t)$$

$$\begin{aligned}\ln(S_{0.5}) &= \ln(40) + \left[0.5\left(0.16 - \frac{0.2^2}{2}\right) + 0.2B_{0.5}\right] \\ &= 3.758879 + 0.2B_{0.5} \sim N(0, t)\end{aligned}$$

那么

$$\frac{\ln(S_{0.5}) - 3.758879}{0.2} = B_{0.5} \sim N(0, 0.5)$$

亦即

$$X = \frac{\ln(S_{0.5}) - 3.758879}{0.2\sqrt{0.5}} \sim N(0, 1)$$

例 2.3.6



练习

- 设 $X \sim N(3, 2^2)$, 求 $P(2 < X \leq 5)$, $P(X > 2)$,
确定 c 使得 $P(X < c) = P(X > c)$.

练习

- 某地区成年男子的体重 X 服从正态分布 $N(\mu, \sigma^2)$, 若已知 $P(X \leq 70) = 0.5$, $P(X \leq 60) = 0.25$.

1. 求 μ, σ ,
2. 若在这个地区随机抽出 5 名成年男子, 问其中至少有两个人的体重超过 65 的概率是多少?

$$Y = \frac{X - \mu}{\sigma} \sim N(0, 1)$$

$$P(X \leq 70) = P\left(Y \leq \frac{70 - \mu}{\sigma}\right) = 0.5$$

$$P(X \leq 60) = P\left(Y \leq \frac{60 - \mu}{\sigma}\right) = 0.25.$$

Outline

2.0 引言

2.1 随机变量

2.2 一维离散型随机变量

2.3 一维连续型随机变量

2.4 一维随机变量的分布

附加

问题

已知 X 的分布, 求 $f(X)$ 的分布

离散型

列举 $f(X)$ 在离散点上的相应取值, 然后将 $f(X)$ 取相同值的概率相加.

例: 设 X 的分布列为

$$\begin{pmatrix} -2 & -1 & 0 & 1 & 2 \\ 0.15 & 0.2 & 0.2 & 0.2 & 0.25 \end{pmatrix}$$

求 $Y = X^2$, $Z = 2X - 1$, $W = X + 1$ 的分布列.

$$\begin{pmatrix} 0 & 1 & 4 \\ 0.2 & 0.4 & 0.4 \end{pmatrix}$$

解

直接求 Y, Z, W 的分布列

$$Y \sim \begin{pmatrix} 4 & 1 & 0 \\ 0.4 & 0.4 & 0.2 \end{pmatrix}$$

$$Z \sim \begin{pmatrix} -5 & -3 & -1 & 1 & 3 \\ 0.15 & 0.2 & 0.2 & 0.2 & 0.25 \end{pmatrix}$$

$$W \sim \begin{pmatrix} 3 & 2 & 1 \\ 0.4 & 0.4 & 0.2 \end{pmatrix}$$

连续型

先求 $f(X)$ 的分布函数, 再对分布函数求导得到 $f(X)$ 的分布密度函数

例: $X \sim N(0, 1)$, 求 $Y = X^2$ 的分布密度函数.

第一步: 先求 Y 的分布函数 $F_Y(y)$. 显然的, 对 $y < 0$ 有

$$P(Y \leq y) = 0$$

对 $y \geq 0$, 有

$$F_Y(y) = P(X^2 \leq y) = P(-\sqrt{y} \leq X < \sqrt{y})$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\sqrt{y}}^{\sqrt{y}} e^{-\frac{x^2}{2}} dx$$

$$\frac{d}{dz} \int_0^z + \int_{-z}^0$$

$$= \frac{2}{\sqrt{2\pi}} \cdot \int_0^{\sqrt{y}} e^{-\frac{x^2}{2}} dx \rightsquigarrow \frac{d}{dy} (\quad) = \sqrt{\frac{2}{\pi}} \cdot e^{-\frac{y}{2}} \cdot \frac{1}{2\sqrt{y}}$$

连续型

$$\left(\frac{1}{\sqrt{2\pi}}\right) \int_{-\sqrt{y}}^{\sqrt{y}} e^{-\frac{x^2}{2}} dx = \dots \int_0^{\sqrt{y}} \dots - \int_0^{-\sqrt{y}} \stackrel{z=\sqrt{y}}{=} \dots \int_0^z - \int_0^{-z} \dots$$

第二步: 对 $F_Y(y)$ 关于 y 求导数, 得到

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y}{2}} \left(\frac{1}{2\sqrt{y}} \right) - \frac{1}{\sqrt{2\pi}} e^{-\frac{y}{2}} \left(-\frac{1}{2\sqrt{y}} \right)$$
$$= \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{y}} e^{-\frac{y}{2}}$$

$$\frac{d}{dy} F_Y(y) = \frac{d}{dz} \left(\int_0^z - \int_0^{-z} \dots \right) \cdot \frac{dz}{dy}$$
$$= \left(e^{-\frac{z^2}{2}} + e^{-\frac{(-z)^2}{2}} \right) \cdot \frac{d\sqrt{y}}{dy}$$

即

$$f_Y(y) = \begin{cases} \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{y}} e^{-\frac{y}{2}}, & y > 0 \\ 0, & y \leq 0 \end{cases}$$

$\frac{1}{\sqrt{2\pi}} \times 2 \cdot e^{-\frac{y}{2}} \cdot \frac{1}{2\sqrt{y}}$

$aX + b$ 的分布

设 X 的分布函数为 f_X , 求 $Y = aX + b$ 的分布密度函数, 其中 $a \neq 0$

记 Y 的分布函数为 F_Y , 分布密度函数为 f_Y , 那么

► 当 $a > 0$ 时

第一步: 求 Y 的分布函数

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(aX + b \leq y) \\ &= P\left(X \leq \frac{y-b}{a}\right) \\ &= \int_{-\infty}^{\frac{y-b}{a}} f_X(x) dx \end{aligned}$$

$aX + b$ 的分布

第二步：对 $F_Y(y)$ 关于 y 求导，有

$$f_Y(y) = \frac{1}{a} f_X\left(\frac{y-b}{a}\right)$$

► 当 $a < 0$ 时，有

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(aX + b \leq y) \\ &= P(X \geq \frac{y-b}{a}) = \int_{\frac{y-b}{a}}^{+\infty} f_X(x) dx \end{aligned}$$

关于 y 求导，有

$$f_Y(y) = \left(-\frac{1}{a}\right) f_X\left(\frac{y-b}{a}\right)$$

$aX + b$ 的分布

综上所述

$$\underline{f_Y(y) = \frac{1}{a} f_X\left(\frac{y-b}{a}\right)}$$

作为特例, 若 $X \sim N(\mu, \sigma^2)$, 且 $Y = \frac{X-\mu}{\sigma}$, 那么由
上式 $= \frac{1}{\sigma}X + (-\frac{\mu}{\sigma})$

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

即 $Y \sim N(0, 1)$, 与命题 2.3.1 吻合.

例子

- ▶ 请大家验证: $X \sim N(\mu, \sigma^2)$, 那么若 $a \neq 0$, 则 $aX + b \sim N(a\mu + b, a^2\sigma^2)$. 并根据这个结论, 说明命题 2.3.1
- ▶ 设随机变量 X 的密度函数是

$$f_X(x) = \begin{cases} \frac{2x}{\pi^2}, & 0 < x < \pi, \\ 0, & \text{其他} \end{cases}$$

求 $Y = \sin(X)$ 的密度函数.