习题 34

- 灭蚊器 10 个：次品 3 个，正品 7 个。
- 售出 2 个后，从剩余中随机抽取 1 个

问：在“抽取的为正品”的条件下，“已售出为一正一次”的概率是多少?

方法一.

解：$B_i = \{$售出 2 个后，剩余正品 i 个$\}$

$P(B_7) = \frac{C_3^2}{C_{10}^2}$，$P(B_6) = \frac{C_3^1 \cdot C_7^1}{C_{10}^2}$，$P(B_5) = \frac{C_7^2}{C_{10}^2}$

$A = \{\cdot \cdots\cdots$，随机取得正品$\}$

$P(A|B_i) = \frac{i}{8}$

$\Rightarrow P(A\cap B_i) = P(B_i)\cdot P(A\cap B_i)$

$\Rightarrow P(B_6|A) = \frac{P(A\cap B_6)}{P(A)} = \frac{P(A\cap B_6)}{\sum_{i=5}^{7} P(A\cap B_i)} = \cdots\cdots$

方法二：不放回取 3 个灭蚊器

$A_i = \{$第 i 次取正品$\}$

要求 $P(A_1\bar{A}_2 \cup \bar{A}_1A_2 | A_3) = \frac{P(A_1\bar{A}_2A_3 \cup \bar{A}_1A_2A_3)}{P(A_3)} = \frac{P(A_1\bar{A}_2A_3)+P(\bar{A}_1A_2A_3)}{P(A_3)}$

习题 36

袋中有球：红 x a；白 x b；黑 x c。
每次有放回地取球 1 个，并再放入同色的球 d 个。
记 $A_k = \{$第 k 次取得红球$\}$。求：$P(A_1|A_k)$。

解：由 35 题得 $P(A_k) = \dfrac{a}{a+b+c}$

取 $k=2$，要求 $P(A_1|A_2) = \dfrac{P(A_1A_2)}{P(A_2)} = \dfrac{a+d}{a+d+b+c}\cdot\dfrac{P(A_1)}{P(A_2)}$

$= \dfrac{a+d}{a+d+b+c}$

$$\boxed{\begin{aligned} P(A_1A_2) &= P(A_2|A_1)\cdot P(A_1) \\ &= \frac{a+d}{a+d+b+c}\cdot P(A_1) \end{aligned}}$$

考察一般的 k.

第.次	1	2	3 ······	k
红	a	$a+d$		?
非红	$b+c$	$b+c$		?

$P(A_k|A_1)$

$= P((k-1)$次取红球$\Big|$红 $= a+d$，非红 $= b+c)$

$= \dfrac{a+d}{a+d+b+c}$

$P(A_1|A_k) = \dfrac{P(A_k|A_1)\cdot P(A_1)}{P(A_k)} = P(A_k|A_1)$

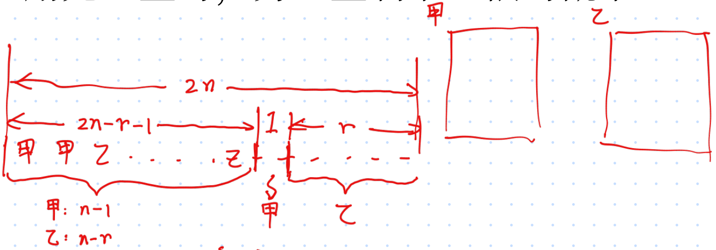
$= \dfrac{a+d}{a+d+b+c}$ #

附加题 2.1

两盒火柴，每盒 n 根。每次随机选一盒火柴，从中用掉一根。

求：用完一盒时，另一盒剩下 r 根的概率？

解：



$$C_{2n-r-1}^{n-1}\times\left(\frac{1}{2}\right)^{2n-r-1}\times\frac{1}{2}\times 2$$

{甲+乙}

(2n-r)th 为甲

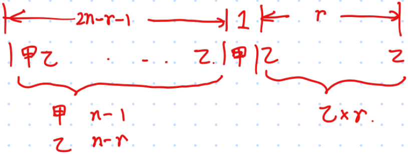
附加题 2.2 ?（①：只考虑前 2n-r ?）

两盒火柴，每盒 n 根。每次随机在所有火柴中选一根用掉

$3\times2\times1$

求：用完一盒时，另一盒剩下 r 根的概率?

解:



$$\frac{C_{2n-r-1}^{n-1}}{C_{2n}^{2n-r-1}} \qquad \frac{2}{r+1}$$

○ ○ . . . ○ $2n$

选n 放甲

$|\Omega| = C_{2n}^{n}$

$A = A_甲 \cup A_乙 = 2\cdot C_{2n-r-1}^{n-1}$

$A_甲 = C_{2n-r-1}^{n-1}$

回顾

- 二次分布
  - n 次独立同分布（二值）试验

    p：成立概率.
  - $X = \{$成立次数$\}$
  - $P(X=k) = C_n^k \cdot p^k \cdot (1-p)^{n-k}$.
- 泊松分布.
  - 单位时间内，事件发生的次数为 λ. e.g. 1 d 卖 3个包子

    $n=24$: 1 h . . $\frac{1}{8}$ . . .
  - $X = \{$单位时间内事件发生的次数$\}$
  - $P_n(X=k) = C_n^k \cdot p_n^k \cdot p_n^{n-k}$

    $\lim\limits_{n\to+\infty} P_n(X=k) = \frac{1}{k!} e^{-\lambda} \cdot \lambda^k$. $(\lambda = \lim n \cdot p_n)$
  - Q: $Y = \{$两次事件之间的时间间隔$\}$

几何分布

定义随机试验，是一次次的做伯努利试验，直到第一次成功为止，记录所做的伯努利试验次数.
设每次独立试验成功的概率为 p，记首次成功的时候做的伯努利试验的次数为 X，那么 $X = k$ 代表事件

$$A = \{\text{第}1, 2, \cdots, k-1\text{次试验都是失败的，} \\ \text{第}k\text{次试验是成功的}\}$$

则有

$$P(A) = P_{\text{geom}}(X = k) = (1-p)^{k-1} p$$

称 X 服从 几何分布，记作 $X \sim Geo(p)$.

例 2.2.6

设一个地下采矿点有 5 个可以上升到地面的通道, 如果事故发生, 只有一个通道可以逃生, 且由于没有照明, 所以遇险者 每次 只能 随意 的在 5 个通道中选择一个, 若发现该通道不通, 则需要返回原点再 随意 的选择一个. 求

- 第三次才选择正确的通道的概率 $(\frac{4}{5})^2 \cdot \frac{1}{5}$.
- 选择成功时已经选择其他错误的通道次数不大于 6 次的概率

$$P(X\le 7) = P(X=1)+P(X=2)+\cdots+P(X=6)+P(X=7)$$

$$= \sum_{i=1}^{7} (1-p)^{i-1}\cdot p^{i}$$

例 2.2.6

解：定义随机变量 X 为选择正确时总共选择的次数, 那么 $X \sim Geo(\frac{1}{5})$

- 第三次才选择正确的概率

$$P_{\mathrm{Geo}}(X=3)=\left(\frac{4}{5}\right)^{2}\frac{1}{5}$$

- 选择正确前选择其他通道的次数不大于 6

$$P_{\mathrm{Geo}}(X\leq 7)=\sum_{k=1}^{7}\left(\frac{4}{5}\right)^{k-1}\frac{1}{5}$$

几何分布的“无记忆性”

几何分布具有“无记忆性”，也就是说当前 k 次试验都不成功时，第 $k+n$ 次试验首次成功与 k 的大小无关，即

$$P(X = k+n \mid X > k) = P(X = n)$$

几何分布的“无记忆性”

证明

$$
\begin{aligned}
P(X=k+n|X>k) &= \frac{P((X=k+n)\cap(X>k))}{P(X>k)} \\
&= \frac{P(X=k+n)}{1-P(X\le k)} \\
&= \frac{(1-p)^{n+k-1}p}{(1-p)^k} \\
&= (1-p)^{n-1}p = P(X=n)
\end{aligned}
$$

作为练习, 请大家验证

$$P(X>k+n|X>k)=P(X>n)$$

课堂练习

袋中有 m 个白球, n 个黑球, 现有放回的摸球, 直到摸到白球停止, 请问已知摸球的次数大于 3, 那么在第 5 次后停止的概率.

$X=\{$拿到白球时, 已经拿球次数$\}$

$P(X=5 \mid X>3) = P(X=2) = \frac{n}{m+n} \cdot \frac{m}{m+n}$

Outline

一维连续型随机变量

若随机变量的可能取值充满 $\mathbb{R}$ 上的一个区间 (a, b), 则称其为 连续随机变量.
请思考下面的问题:

- 计算质量密度为 $\rho(x)$ 的一维杆的一个端点到杆上某个指定点的质量
- 在 $[a, b]$ 区间随机的投点, 求落点位置 X 的分布函数

一维连续型随机变量的定义

定义 2.3.1

设 $(\Omega, \mathcal{F}, P)$ 为概率空间, X 为其上的随机变量, F_x 为 X 的分布函数. 如果存在 非负可积 的函数 f_X, 使得

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt, \quad \forall x \in (-\infty, \infty)$$

那么称 X 为 连续型随机变量, 称 $f_X(x)$ 为 X 的 分布密度函数.

分布密度函数与分布函数

由 Newton-Leibniz 公式, 在 $f_X(x)$ 的连续点上, 有

$$f_X(x) = F_X'(x) = \frac{d}{dx}\int_{-\infty}^{x} f_X(t)\,dt = f_X(x) - \underbrace{f_X(-\infty)}_{=0}$$

- 若 $f_X(x)$ 是 X 的分布密度函数, 那么
  - $f_X(x) \geq 0, \ \forall x \in (-\infty, \infty)$
  - $\int_{-\infty}^{\infty} f_X(x)dx = 1$
- 若定义在 $(-\infty, \infty)$ 上的函数 f 满足上面的两点, 那么令

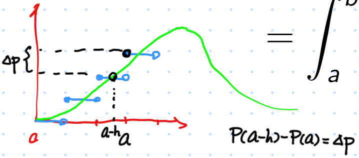
$$F(x) = \int_{-\infty}^{x} f(t)dt, x \in (-\infty, \infty)$$

那么 $F(x)$ 一定是某随机变量的分布函数.

分布密度函数与分布函数

显然的

$$P(a < X \le b) = F_X(b) - F_X(a) = \int_a^b f_X(x)dx, \ \forall a < b \in \mathbb{R}$$



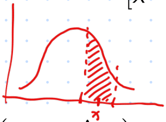
而

$$P(X = a) = \lim_{h \to 0^+} P(a - h < X \le a) = \lim_{h \to 0^+} \int_{a-h}^a f_X(x)dx = 0$$

上面的式子表明连续型随机变量取任意的单点值的概率为零，换句话说，连续型随机变量的分布特性不能通过列举它取单点值的概率表示.

分布密度函数与分布函数

根据 Riemann 积分的性质, 存在 $\eta \in (m, M)$, 其中 $m = \min_{[x-\Delta x, x+\Delta x]} f_X(x)$, $M = \max_{[x-\Delta x, x+\Delta x]} f_X(x)$ 使得

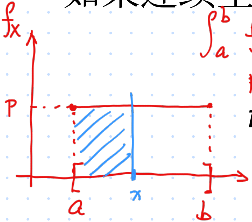


$$
\begin{aligned}
\eta \Delta x &= \int_{x-\Delta x}^{x+\Delta x} f_X(t) t = F_X(x+\Delta x) - F_X(x-\Delta x) \\
&= P(x-\Delta x < X \le x+\Delta x)
\end{aligned}
$$

说明连续型随机变量在密度函数取值大的点附近取值的概率也较大.

均匀分布

如果连续型随机变量 X 的分布密度函数为



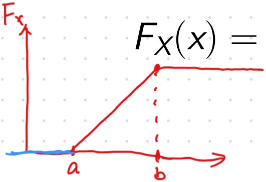
$\int_a^b f_X(t)dt = 1$

$p(b-a) = 1$

$$f_X(x) = \begin{cases} \dfrac{1}{b-a}, & \text{当} x \in [a, b] \\ 0, & \text{当} x \notin [a, b] \end{cases}$$

则称 X 服从 $[a, b]$ 上的 均匀分布, 记为 $X \sim U[a, b]$, 其分布函数为

uniform



$$F_X(x) = \int_{-\infty}^{x} f(t)t = \begin{cases} 0, & \text{当} x < a \\ \dfrac{x-a}{b-a}, & \text{当} a \leq x \leq b \\ 1, & \text{当} x > b \end{cases}$$

课堂练习

- 设随机变量 $X \sim U[0,10]$, 对 X 进行 4 次独立观测, 求至少有 3 次观测值大于 5 的概率.
- 在 $[0,1]$ 上任取一点记为 X, 求 $P(X^2 - \frac{3}{4}X + \frac{1}{8} \geq 0)$

1. 解: $A = \{$观测值$>5\} = \frac{1}{2}.$

$P(X>5) = 1 - P(X \leq 5)$

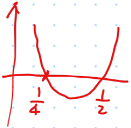
$= 1 - F_X(5) = \frac{10-5}{10-0} = \frac{1}{2}.$

$Y = \{A$ 发生的次数$\}$

$P(Y \geq 3) = P(Y=3) + P(Y=4)$

$= C_4^3 \cdot P(A)^3 \cdot P(\bar{A})^1 + C_4^4 \cdot P(A)^4 \cdot P(\bar{A})^0$

2. $X^2 - \frac{3}{4}X + \frac{1}{8} \geq 0$ 。



$\Leftrightarrow (X - \frac{1}{4})(X - \frac{1}{2}) \geq 0$

$\Leftrightarrow \{X \leq \frac{1}{4}\} \cup \{X \geq \frac{1}{2}\}$

$\Leftrightarrow X \sim U[0,1].$

求 $P(X \leq \frac{1}{4}) + P(X \geq \frac{1}{2}) = \frac{3}{4}$

#