

习题课 2

罗雁天

March 25, 2019

1. 再次考虑“匹配”问题。 n 个人随机地选取帽子，试问恰好戴上了自己帽子的人数的期望和方差是多少。

解. 带上自己帽子的人数 X 满足:

$$X = H_1 + H_2 + \cdots + H_n, \quad H_k = \begin{cases} 1 & \text{第 } k \text{ 个人戴对帽子} \\ 0 & \text{第 } k \text{ 个人戴错帽子} \end{cases} \quad (1)$$

则由

$$P(H_k = 1) = 1 - P(H_k = 0) = \frac{1}{n} \quad (2)$$

得到

$$E(H_k) = 1 \times P(H_k = 1) + 0 \times P(H_k = 0) = \frac{1}{n} \quad (3)$$

从而

$$\begin{aligned} E(X) &= E(H_1 + H_2 + \cdots + H_n) \\ &= E(H_1) + E(H_2) + \cdots + E(H_n) \\ &= n \times \frac{1}{n} \\ &= 1 \end{aligned} \quad (4)$$

即戴上自己帽子的人数的期望值为 1。也就是说，平均只有一个人戴上自己的帽子。相比于上次习题课得到的人数分布的结果，这里给出的期望能够更加直观地表明恰好匹配（戴上自己的帽子）是较为困难的事情。

接下来求方差，首先求 $E(X^2)$ 如下：

$$\begin{aligned} E(X^2) &= E((H_1 + H_2 + \cdots + H_n)^2) \\ &= E\left(\sum_{k=1}^n H_k^2 + \sum_{i=1}^n \sum_{j \neq i}^n H_i H_j\right) \\ &= \sum_{k=1}^n E(H_k^2) + \sum_{i=1}^n \sum_{j \neq i}^n E(H_i H_j) \end{aligned} \quad (5)$$

根据我们的定义，我们可以得到：

$$E(H_k^2) = 1 \times P(H_k = 1) + 0 \times P(H_k = 0) = \frac{1}{n} \quad (6)$$

$H_i H_j$ 表示第 i, j 两人均带对帽子，因此：

$$P(H_i H_j = 1) = 1 - P(H_i H_j = 0) = \frac{1}{n(n-1)} \quad (7)$$

因此：

$$E(H_i H_j) = 1 \times P(H_i H_j = 1) + 0 \times P(H_i H_j = 0) = \frac{1}{n(n-1)} \quad (8)$$

带入 5 得：

$$E(X^2) = n \times \frac{1}{n} + n(n-1) \times \frac{1}{n(n-1)} = 2 \quad (9)$$

因此：

$$Var(X) = E(X^2) - [E(X)]^2 = 1 \quad (10)$$

□

2. n 对夫妇共 $2n$ 人呆在一个房间内，现从中随机挑选 m 个人走出房间，问房间内还剩下的未被拆散的夫妇数目的均值。

解. 房间内剩下的未被拆散的夫妇数目 X 满足

$$X = C_1 + C_2 + \cdots + C_n, \quad C_k = \begin{cases} 1 & \text{第 } k \text{ 对夫妇在房间} \\ 0 & \text{第 } k \text{ 对夫妇不在房间} \end{cases} \quad (11)$$

则由

$$P(C_k = 1) = 1 - P(C_k = 0) = \frac{\binom{2n-2}{m}}{\binom{2n}{m}} = \left(1 - \frac{m}{2n}\right) \left(1 - \frac{m}{2n-1}\right) \quad (12)$$

得到

$$E(C_k) = 1 \times P(C_k = 1) + 0 \times P(C_k = 0) = \left(1 - \frac{m}{2n}\right) \left(1 - \frac{m}{2n-1}\right) \quad (13)$$

从而

$$\begin{aligned} E(X) &= E(C_1 + C_2 + \cdots + C_n) \\ &= E(C_1) + E(C_2) + \cdots + E(C_n) \\ &= n \times \left(1 - \frac{m}{2n}\right) \left(1 - \frac{m}{2n-1}\right) \end{aligned} \quad (14)$$

□

3. 某商场发行 n 种购物券，每一次在该商场购物，即可获得一张种类随机选取的购物券。该商场规定，如果能够收集齐所有的购物券，那么就可以得到该商场的大奖。问获得大奖所需要的购物次数的期望是多少。

解. 收集齐所有的购物券所需要的购物次数 X 可以写为

$$X = C_1 + C_2 + \cdots + C_n \quad (15)$$

其中 C_k 表示在收集到 $k-1$ 种购物券的前提下, 收集到第 k 种购物券所需要的购物次数。由于手中已经有 $k-1$ 种购物券, 因此收集到新购物券的可能只有 $n-k+1$ 种, 所以 C_k 是几何分布随机变量, 参数为 $p = (n-k+1)/n$, 因此我们有:

$$E(C_k) = \frac{1}{p} = \frac{n}{n-k+1} \quad (16)$$

所以:

$$\begin{aligned} E(X) &= E(C_1 + C_2 + \cdots + C_n) \\ &= E(C_1) + E(C_2) + \cdots + E(C_n) \\ &= 1 + \frac{n}{n-1} + \frac{n}{n-2} + \cdots + \frac{n}{1} \\ &= n \left(1 + \frac{1}{2} + \cdots + \frac{1}{n} \right) \end{aligned} \quad (17)$$

由微积分的知识可以知道 $E(X) \approx n \ln n$, 可见, 当 n 较大时, 如果想收集到所有的购物券, 则需要付出数倍于购物券数目的购物次数 (以 30 张购物券为例, $\ln(30) \approx 3.4$)。这正是商家的精明之处。 \square

4. 设 R 罐中有 n 个红球, H 罐中有 n 个黑球。每次操作都从两个罐中各随机取出一球, 交换放置 (R 罐中取出的放入 H 罐, H 罐中取出的放入 R 罐), 问 k 次操作后, R 罐中的红球数目的均值。

解. k 次操作后 R 罐的红球数目满足:

$$X = R_1 + R_2 + \cdots + R_n, \quad R_t = \begin{cases} 1 & \text{第 } t \text{ 个红球仍在 } R \text{ 罐} \\ 0 & \text{第 } t \text{ 个红球不在 } R \text{ 罐} \end{cases} \quad (18)$$

对于操作开始前 R 罐中原有的每一个红球而言, 其经过 k 次操作仍回到 R 罐意味着被选中的次数为偶数 (从来未被选中意味着选中次数为 0, 仍为偶数)。所以:

$$P(R_t = 1) = 1 - P(R_t = 0) = \sum_{m=2i} \binom{k}{m} \left(\frac{1}{n} \right)^m \left(1 - \frac{1}{n} \right)^{k-m} \quad (19)$$

根据二项式定理, 我们可以计算出偶数次项的和可以计算如下:

$$\sum_{m=2i} \binom{k}{m} a^m b^{k-m} = \frac{1}{2} \left((b+a)^k + (b-a)^k \right) \quad (20)$$

所以:

$$\begin{aligned} P(R_t = 1) &= \frac{1}{2} \left(\left(1 - \frac{1}{n} + \frac{1}{n} \right)^k + \left(1 - \frac{1}{n} - \frac{1}{n} \right)^k \right) \\ &= \frac{1}{2} \left(1 + \left(1 - \frac{2}{n} \right)^k \right) \end{aligned} \quad (21)$$

从而

$$\begin{aligned}
 E(X) &= E(R_1 + R_2 + \cdots + R_n) \\
 &= E(R_1) + E(R_2) + \cdots + E(R_n) \\
 &= \frac{n}{2} \left(1 + \left(1 - \frac{2}{n} \right)^k \right)
 \end{aligned} \tag{22}$$

□

5. 设随机变量 $X \sim N(0, \sigma^2)$, $Y = [X]$, 即 Y 是 X 的整数部分 (例如: $[1.2] = 1, [-2.3] = -3$), 计算 Y 的期望

解. 设 X 的分布函数为 $f_X(x)$, Y 是整数集上的离散随机变量, 我们有:

$$P(Y = n) = \int_n^{n+1} f_X(x) dx \tag{23}$$

考虑到 $f_X(x)$ 是偶函数, 所以:

$$\begin{aligned}
 E(Y) &= \sum_{n=-\infty}^{+\infty} n \times P(y = n) \\
 &= \sum_{n=-\infty}^{+\infty} n \times \int_n^{n+1} f_X(x) dx \\
 &= \int_1^2 f_X(x) dx + 2 \int_2^3 f_X(x) dx + \cdots \\
 &\quad - \int_0^1 f_X(x) dx - 2 \int_1^2 f_X(x) dx - 3 \int_2^3 f_X(x) dx - \cdots \\
 &= - \int_0^{+\infty} f_X(x) dx \\
 &= -\frac{1}{2}
 \end{aligned} \tag{24}$$

□

6. 设有 k 种不同的优惠券, 每次收集到第 i 种优惠券的概率为 p_i , $\sum_{i=1}^k p_i = 1$, 且每次收集之间是相互独立的。如果收集了 n 张优惠券, 那么优惠券的种类的期望是多少?

解. 优惠券的种类数 X 满足:

$$X = H_1 + H_2 + \cdots + H_k, \quad H_i = \begin{cases} 1 & \text{收集到第 } i \text{ 种优惠券} \\ 0 & \text{没有收集到第 } i \text{ 种优惠券} \end{cases} \tag{25}$$

那么我们有:

$$P(H_i = 1) = 1 - P(H_i = 0) = 1 - (1 - p_i)^n \tag{26}$$

所以：

$$E(H_i) = 1 \times P(H_i = 1) + 0 \times P(H_i = 0) = 1 - (1 - p_i)^n \quad (27)$$

从而：

$$\begin{aligned} E(X) &= E(H_1 + H_2 + \cdots + H_k) \\ &= E(H_1) + E(H_2) + \cdots + E(H_k) \\ &= \sum_{i=1}^k (1 - (1 - p_i)^n) \end{aligned} \quad (28)$$

□

7. 在单位圆上随机取两个点构成一条弦，试计算从原点到该弦的距离所服从的概率密度函数

解. 设两点对圆心的张角为 θ ，则 $\theta \sim U(0, \pi)$ ，我们有：

$$f_\theta(\theta) = \frac{1}{\pi} \quad (29)$$

设从原点到弦的距离为 r ，则 $r = \cos\left(\frac{\theta}{2}\right)$ 。由于在 $\theta \in [0, \pi], r \in [0, 1]$ 时， r 是 θ 的单调函数，于是：

$$f_r(r) = f_\theta(\theta) \left| \frac{d\theta}{dr} \right| = \frac{2}{\pi\sqrt{1-r^2}}, r \in [0, 1] \quad (30)$$

注意：这里的随机取点认为是两者所对的圆心角（取范围 $[0, \pi]$ ）服从均匀分布，若选用其他方式，例如弦中点到圆心的距离服从均匀分布，则会得到不同的结果，这与 Bertrand 悖论的道理是相同的，即不同的样本空间会导致不同的结果。

□