第一周作业

助教-夏小凡

2024 年 3 月 6 日

求期望的两个常见思路一个是利用期望的线性性，一个是利用条件期望，以下讲的两个方法无论考试还是作业还是学科内都极度常见

线性性：若 $Y = X_1 + \cdots + X_n$，则

$$E[Y] = E[X_1] + \cdots + E[X_n]$$

取条件期望 (离散情况的求和就是有限求和或者可数求和，连续情况的求和就是积分):

$$E[Y] = E[E[Y|X]] = \sum_x E[Y|X = x]P(X = x)$$

此时可先求出 $E[Y|X = x]$，再对它取期望。

至于求方差的话，类似的有

$$Var[Y] = \sum_{i=1}^{n}\sum_{j=1}^{n} Cov(X_i, X_j)$$

$$Cov(X, Y) = Cov(E[X|Z], E[Y|Z]) + E[Cov(X, Y|Z)]$$

第二个式子 let $X = Y$ 即可获得利用条件期望求方差的公式。

1 作业题解答

题目 1: 课本第一章的习题 6

解答:(a):

记 I_j 为 j 时刻产生的记录个数，那么 I_j 的取值显然是 0 或者 1，并且可以发现:

$$P(I_j=1)=P(X_j\text{是前 j 个随机变量里最大的})$$

对称性告诉我们前 j 个随机变量是最大的概率都是一样的。因此有 $P(I_j=1)=\frac{1}{j}$

从而有

$$E[N_n]=\sum_{j=1}^{n}E[I_j]=\sum_{j=1}^{n}\frac{1}{j}$$

至于求方差，事实上容易发现 I_j 之间是相互独立的。(第 n+1 个是否是最大的显然不取决于前面 n 个里面哪个最大)，因此我们有

$$Var[N_n]=\sum_{j=1}^{n}Var(I_j)=\sum_{j=1}^{n}\frac{1}{j}(1-\frac{1}{j})$$

(b):

$$P(T>n)=P(I_2=\cdots=I_n=0)=\Pi_{j=2}^{n}(1-\frac{1}{j})=\frac{1}{n}$$

所以有

$$P(T<\infty)=\lim_{n\to\infty}P(T\le n)=1$$

$$E[T]=\sum_{n=0}^{\infty}P(T>n)=\infty$$

(c):

这里考虑利用 $P(A|B)=P(A)$ 来证明独立性。

$$P(X_{T_y}>x|T_y=n)$$

$$=P(X_n>x|X_1\le y,\cdots,X_{n-1}\le y,X_n>max(X_1,\cdots,X_{n-1}),X_n>y)$$

注意到 X_n 是前 n 个里面最大的这一条件可被其余条件推出，所以可以忽略这个条件，又注意到 X_i 的独立性让我们在已知 $X_i, i < n$ 的时候无法推出 X_n 的信息, 所以上述条件本质可以简化为

$$P(X_n > x | X_n > y)$$

这得出的自然会是和 n 无关的数字，也就是和 T_y 的取值无关，所以独立

题目 2: 课本第一章的习题 7

解答:这题就是常见的超几何分布，分布列大家肯定都会写，直接硬算期望和方差需要不少组合数的技巧，我反正记不住，分享一下不硬算的方法:

不妨考虑给 m+n 个物体编号，其中白球的编码是 1–n，然后我们抽取 k 次对应 k 个随机变量 X_i,$X_i = 1$ 表示第 i 次抽到了白球（X_i 是第 i 次抽到的白球个数）。

我们不难发现 (第 i 次抽到 j 号球的概率，可以考虑给 m+n 个球进行排列，排列的前 k 个是我们抽中的，然后第 i 次抽到 j 号的总排列数就是固定 i 位置为 j 号球对别的地方全排列)

$$P(X_i = 1) = \sum_{j=1}^{n} P(\text{第 i 次抽到 j 号球}) = \frac{n}{m+n}$$

$$P(X_i = 1, X_j = 1) = C_n^2 * A_2^2 * \frac{1}{n + m(n+m-1)} = \frac{n(n-1)}{(m+n)(m+n-1)}$$

从而我们可以知道 ($E[X_i], E[X_i^2], E[X_iX_j]$ 都已经算出来了，剩下的就是自己慢慢求和):

$$E[X] = E[X_1 + \cdots + X_k] = \frac{kn}{m+n}$$

$$Var[X] = \frac{kmn(m+n-k)}{(m+n)^2(m+n-1)}$$

题目 3: 课本第一章的习题 12

解答:题目条件意味着 (几乎处处有)

$$(a - X)X \geq 0$$

两边取期望可知

$$aE[X] \geq E[X^2]$$

从而我们有

$$E[X^2] - E[X]^2 \leq aE[X] - E[X^2] \leq \frac{a^2}{4}$$

题目 3 的注记:这是概率论矩估计的一个常见技巧: 先构建一个关于 X 的不等式，然后两边取期望即可得到一个关于 E[X] 的不等式。矩估计的另一个有趣例子放在补充习题里面了。

题目 4: 课本第一章的习题 14（题目意思是，反复投骰子直到累计投出了 10 次偶数停止）

解答:这就是一个负二项分布的模型 (几何分布的自然推广)，因此可以考虑使用几何分布的性质来研究，也可以很容易写出对应的概率密度函数。[当然求期望方差什么的还是按照我们之前说的，取条件或者找拆分，硬算很慢而且容易算错，只是个理论可行但实际上很难可行的做法]

(a): 我们考虑熟知的几何分布（进行 k 次伯努利实验才得到第一次成功的的模型）的结论：如果成功概率为 p, 那么实验次数 k 的期望是 $\frac{1}{p}$, 方差是 $\frac{1-p}{p^2}$

此时对于每次实验，我们可以认为投出偶数叫做成功，投出奇数叫做失败，也就是 $p = \frac{1}{2}$, 那么直到第一次投出偶数的时候，投掷的总次数期望应该是 2, 投掷的总偶数次数期望显然是 1，因此投掷的总奇数期望就是 2-1=1。对称性原理 $[E[X_1] = E[X_3] = E[X_5]]$ 告诉我们投掷的总 1 数期望应该是 $\frac{1}{3}$，然后上述过程重复 10 次，也就是说 $E[X_1] = \frac{10}{3}$

(b): 对称性告诉我们

$$E[X_2] = E[X_4] = E[X_6]$$

投掷结束的时候显然有

$$E[X_2] + E[X_4] + E[X_6] = 10$$

从而显然

$$E[X_2]=\frac{10}{3}$$

(c): 此时我们可以认为投出偶数叫做成功，投出 1 叫做失败，如果投出 3 或者 5，由于他既不影响投出的总偶数次数，也不影响 X_1，因此我们可以认为投出 3 和 5 就当啥都没有发生，接着重新投骰子进行实验，因此这对应一个 $p=\frac{3}{4}$ 的超几何分布模型。我们有

$$P(\text{总次数为 n})=C_{n-1}^{9}(1-p)^{n-10}p^{10}$$

注意到总次数为 n 的时候相当于投出了 n-10 次 1，从而我们有

$$P(X_1\text{=n})=P(\text{总次数为 n+10})=C_{n+9}^{9}(1-p)^{n}p^{10}$$

(d): 此时也就是总共有 10 次投出偶数的情况，每次都等可能的会是 2,4,6, 实际上出现 2 的次数 X_2 就是一个单纯的满足总实验次数为 10，$p=1/3$ 的二项分布，概率质量函数显然就是

$$P(X_2=n)=C_{10}^{n}p^{n}(1-p)^{10-n}$$

题目 4 的注记:这里是注记.

题目 5: 课本第一章的习题 16

解答:题目的大概意思就是，用计算机按照密度函数为 g 的分布随机生成一个数字 Y，按照均匀分布随机生成一个数字 U, 如果满足 $U\le\frac{f(Y)}{cg(Y)}$，那么就输出一个数字 X, 他的数值和之前随机生成的 Y 一样，如果不满足这个条件，那么就重新生成 Y 和 U，第一问想当于让你求如此产生的 X 服从怎样的分布

题目条件事实上告诉我们

$$P(X=x)=P(Y=x|U\le\frac{f(Y)}{cg(Y)})$$

我们考虑化简右边的式子，他自然可以写成

$$\frac{P(Y=x,U\le\frac{f(Y)}{cg(Y)})}{P(U\le\frac{f(Y)}{cg(Y)})}=\frac{P(Y=x)P(U\le\frac{f(Y)}{cg(Y)}|Y=x)}{P(U\le\frac{f(Y)}{cg(Y)})}$$

注意到 $P(U \leq x) = x$，分子可以被很容易计算:

$$P(Y = x) = g(x), P(U \leq \frac{f(Y)}{cg(Y)}|Y = x) = P(U \leq \frac{f(x)}{cg(x)}) = \frac{f(x)}{cg(x)}$$

故分子也就是 $\frac{f(x)}{c}$

至于分母，我们可以通过计算分子的时候发现，分母在取条件后非常好计算，因此我们对分母取条件就有

$$P(U \leq \frac{f(Y)}{cg(Y)}) = \sum_x P(U \leq \frac{f(Y)}{cg(Y)}|Y = x)P(Y = x)$$

$$= \int_R \frac{f(x)}{cg(x)} g(x)dx = \frac{1}{c}$$

所以也就有 $P(X = x) = f(x)$

至于第二问，我们刚刚计算的分母其实就是被接受的概率，他是 $\frac{1}{c}$，所以迭代次数当然就是均值为 c 的几何随机变量。(他显然是几何随机变量，概率是 1/c 意味着均值是 c)

题目 5 的注记:这题其实一个是理解条件概率的现实意义 (在给定某个条件的情况下，事件发生的概率)，以及条件概率公式的应用

题目 6: 课本第一章的习题 21

解答:题目意思大概就是说一堆在 (0,1) 随机分布的数 U_i 乘起来，如果 n+1 个数字乘起来的时候这个乘积首次小于 $e^{-\lambda}$, 那么 N=n, 我们要研究 $P(N = n)$

n=0 的时候，我们有

$$P(N = 0) = U_1 \leq e^{-\lambda} = e^{-\lambda}$$

所以结论成立，现在假设结论对 n 成立，我们证明结论对 n+1 成立。

$$P(N = n + 1) = P(\Pi_{i=1}^{n+1} U_i \geq e^{-\lambda} > \Pi_{i=1}^{n+2} U_i)$$

$$= \sum_x P(\Pi_{i=1}^{n+1} U_i \geq e^{-\lambda} > \Pi_{i=1}^{n+2} U_i | U_1 = x)P(U_1 = x)$$

$$= \sum_x P(\Pi_{i=2}^{n+1} U_i \geq \frac{e^{-\lambda}}{x} > \Pi_{i=2}^{n+2} U_i) P(U_1 = x)$$

求和式子的乘积第一项显然只当 $\frac{e^{-\lambda}}{x} \in (0,1)$ 的时候不为 0，也就是说我们的求和只需要遍历 x 从 $e^{-\lambda}$ 到 1，并且此时根据归纳假设有

$$P(\Pi_{i=2}^{n+1} U_i \geq \frac{e^{-\lambda}}{x} > \Pi_{i=2}^{n+2} U_i) = \frac{e^{-\lambda}(\lambda + lnx)^n}{xn!}$$

把求和写为积分形式也就有（$P(U_1 = x) = f_{U_1}(x)dx = dx$）

$$= \int_{e^{-\lambda}}^{1} \frac{e^{-\lambda}(\lambda + lnx)^n}{xn!} dx$$

这个积分的结果是 $\frac{e^{-\lambda}\lambda^{n+1}}{(n+1)!}$ 非常显然，也就证完了。

题目 6 的注记:看清楚条件概率公式咋么用了嘛

题目 7: 课本第一章的习题 25

解答:根据提示，我们有

$$M_i = \frac{1}{2}(M_{i-1} + 1) + \frac{1}{2}(M_{i+1} + 1)$$

也就是

$$M_i = 1 + 0.5M_{i-1} + 0.5M_{i+1}$$

$$M_{i+1} - M_i = M_i - M_{i-1} - 2$$

边界条件为 $M_0 = M_k = 0$

也就是说 $M_{i+1} - M_i$ 是公差为-2 的等差数列，所以有

$$M_{i+1} - M_i = (M_1 - M_0) - 2i = M_1 - 2i$$

我们对上面的式子，i 从 0 取到 k-1 进行求和，可以得到

$$0 = M_k - M_0 = kM_1 + k - k^2$$

从而解得 $M_1 = k - 1$, 也就是说 $M_{i+1} - M_i = k - 1 - 2i$

对上述式子 i 从 0 到 j-1 求和即可知道会有

$$M_j = j(k - j)$$

题目 7 的注记:这是学过概率论的人应该掌握的东西：把事件分解后考虑利用递推公式求概率或期望。

题目 8: 课本第一章的习题 30

解答:这题我建议你先做了 31 题再过来，因为我要用那边的结论。

记 A:A 最后被服务，B:B 比 C 早，那么有

$$P(A) = P(A|B)P(B) + P(A|B^c)P(B^c)$$

根据 31 题的结论我们有

$$P(B) = \frac{\lambda_1}{\lambda_1+\lambda_2}, P(B^c) = \frac{\lambda_2}{\lambda_1+\lambda_2}$$

根据无记忆性（B 和 C 后走的那个人，已经等待的时间等于白等了）我们又有

$$P(A|B) = P(B^c), P(A|B^c) = P(B)$$

代入可以算出答案是

$$\frac{2\lambda_1\lambda_2}{(\lambda_1+\lambda_2)^2}$$

题目 9: 课本第一章的习题 31

解答:求最大值最小值分布一律都是用分布函数易于计算的特点，求条件分布硬算即可

$$P(Z > z) = P(X > z)P(Y > z) = e^{-(\lambda_1+\lambda_2)z}$$

[这个相当于告诉你一个结论：两个速率为 v_1, v_2 的泊松过程的叠加过程的速率是 $v_1 + v_2$]

$$P(Z \le z|Z = X) = P(X \le z|Y > X) = \frac{P(X \le z, Y > X)}{P(Y > X)}$$

分子其实就是计算 $P(X < min(z, Y))$, 也就是积分

$$\int_{-\infty}^{z} P(X < y)f_Y(y)dy + \int_{z}^{\infty} P(X \le z)f_Y(y)dy$$

$$= \int_0^z (1-e^{-\lambda_1 y})\lambda_2 e^{-\lambda_2 y}dy + \int_z^\infty (1-e^{-\lambda_1 z})\lambda_2 e^{-\lambda_2 y}dy$$

总之慢慢算会发现答案就是

$$\frac{\lambda_1}{\lambda_1+\lambda_2}(1-e^{-(\lambda_1+\lambda_2)z})$$

至于 $P(Y > X)$

$$\int_0^{+\infty} P(X<y)f_Y(y)dy = \int_0^{+\infty}(1-e^{-\lambda_1 y})\lambda_2 e^{-\lambda_2 y}dy = \frac{\lambda_1}{\lambda_1+\lambda_@}$$

题目 10: 课本第一章的习题 39

解答:和 25 几乎完全一样的思路，我就写简单一点吧。首先递推公式和边界条件分别是

$$E[T_n]=0, E[T_1]+1=E[T_0], E[T_i]=\frac{1}{2}(E[T_{i+1}]+E[T_{i-1}])+1$$

我们不难推出

$$E[T_0]=E[T_2]+4$$

再用归纳法就可以推出来

$$E[T_0]=E[T_i]+i^2$$

从而 $E[T_0]=E[T_n]+n^2=n^2$ 当然你像我 25 题那样慢慢解也可以

题目 11: 课本第二章的习题 4

解答:(拆分是因为没独立性压根算不了)

$$E[N(t)N(t+s)] = E[N(t)N(t)]+E[N(t)(N(t+s)-N(t))] = E[N^2(t)]+E[N(t)]E[N(s)] = \lambda^2t^2+\lambda t$$

[我想泊松分布的均值方差应该是概率论中太常见的东西了，记忆法也可以，或者背个特征函数自己算]