概率统计 B _ 2020-2021 年春季学期期中考试参考答案

2021 年 4 月 22 日

(1) 假设某一件产品的使用寿命（单位为年）X 服从参数 $\lambda = 1$ 的指数分布

- (5 pt) 试求该产品在 3-4 年间损坏（到达寿命）的概率。

  Solution: $\lambda = 1$ 的指数分布概率密度函数为 $p(x) = e^{-x}$。因此 $P(3 < X < 4) = \int_3^4 e^{-x}dx = e^{-3} - e^{-4}$。

- (5 pt) 试写出随机变量 $Y = 2X + 1$ 的概率密度函数

  Solution: X 的取值范围为 $(0, \infty)$，因此 $Y = 2X + 1$ 取值范围为 $(1, \infty)$。对于一切 $y > 1$

  $$P(Y \le y) = P(2X + 1 \le y) = P(x \le \frac{y-1}{2}) = \int_0^{\frac{y-1}{2}} e^{-x}dx$$

  对于 y 求导我们有

  $$p_Y(y) = \begin{cases} \frac{1}{2}e^{-(y-1)/2} & y > 1 \\ 0 & y \le 1 \end{cases}$$

(2) (15 pt) 假设你在燕南园看到橘猫的数目服从参数 $\lambda = 10$ 的泊松分布。已知十个橘猫九个胖，请问你看到的“胖橘猫”的数目服从什么分布?

Solution: 记胖橘猫的数目为 X 要求概率质量函数 $P(X = k)$, 根据全概率公式

$$\begin{aligned} P(X = k) &= \sum_{n=0}^{\infty} P(\text{有 k+n 个橘猫})P(\text{有 k 个胖橘猫}|\text{有 k+n 个橘猫}) \\ &= \sum_{n=0}^{\infty} \frac{10^{k+n}}{(k+n)!}e^{-10} \times C_{n+k}^{k}(9/10)^k(1/10)^n \\ &= \frac{9^k}{k!}e^{-10} \times \left(\sum_{n=0}^{\infty} \frac{1}{n!}\right) = \frac{9^k}{k!}e^{-9} \end{aligned}$$

“胖橘猫”的数目服从参数 $\lambda = 9$ 的泊松分布

(3) 某款电脑游戏玩家每一局均有 1/3 的概率升级到下一关，1/3 的概率留在本关，1/3 的概率被击败出局，每局结果相互独立。假设玩家目前位于第一关。

- (10 pt) 令随机变量 X 为该玩家到出局时所经过的局数。求 X 的期望。

Solution: 推广上一问的结果，我们有一般公式 $P(X=n)=(2/3)^{n-1}\times 1/3$。因此期望为:

$$\begin{aligned}\sum_{n=1}^{\infty} n(2/3)^{n-1}\times 1/3 &= \left(\sum_{n=1}^{\infty}(2/3)^{n-1}\times 1/3\right)+\left(\sum_{n=2}^{\infty}(2/3)^{n-1}\times 1/3\right)+\cdots \\ &= \sum_{n=0}^{\infty}(2/3)^{n}=3.\end{aligned}$$

- (10 pt) 求该玩家在出局前成功升级到第 4 关的概率。

  Solution: 令 p_1, p_2, p_3 分别为假设玩家目前位于第 1，2，3 关时，其在出局前成功升级到第 4 关的概率。由全概率公式，我们有

$$\begin{aligned}p_1 &= p_1/3+p_2/3\\ p_2 &= p_2/3+p_3/3\\ p_3 &= p_3/3+1/3\end{aligned}$$

  解得 $p_1=1/8$。

(4) (15 pt) 本班共有同学 121 人，假设闰年不存在，每个人每天过生日的概率相等且互相独立。请估算至少两位同学生日相同的概率。(提示：可以使用近似 $\exp(-x)\approx 1-x$)

Solution: 这里我们计算关心的概率的反面——即全体同学生日不同的概率。将全体同学任意排成一列。首先记录第一位同学的生日。第二位同学与第一位生日不同的概率为 $(1-1/365)$。在已知前两位同学生日不同的情况下，第三位同学生日与他们都不同的条件概率是 $(1-2/365)$。同理，已知前 n 位同学生日不同的情况下，第 $n+1$ 位同学生日与他们都不同的条件概率是 $(1-n/365)$. 因此，我们 121 人生日都不同的概率为

$$p=(1-1/365)\times(1-2/365)\times\cdots\times(1-120/365)$$

根据提示 $p\approx\exp(\frac{-\sum_{n=1}^{120}n}{365})=\exp(-19.56)=3*10^{-8}\approx 0$。因此 $1-p\approx 1$ 即至少两位同学生日相同的概率几乎为 1。

(5) 假设某种产品的强度服从 $\mu=50, \sigma=2$ 的高斯分布。强度不超过 46 者为次品。

- (5 pt) 求该产品的次品率。

  Solution: 根据高斯分布线性变换，次品率等于 $P(N(0,1)\leq -2)=1-P(N(0,1)\leq 2)$。查表得到次品率 $\approx 2.28\%$

- (10 pt) 客户订购了一批共 100 件此种产品；假设生产商也恰好发货 100 件。但由于次品率的存在，客户收到的正品数目可能不足 100 件。一旦上述情况发生，客户将对生产商给出差评。请问客户给出差评的概率有多大?

  Solution: 此时次品数目近似服从 $\lambda=2.28$ 的泊松分布。故无次品的概率约为 $\exp(-2.28)\approx 0.1$。即客户给出差评的概率高达 90%

- (10 pt) 在现实中，生产商此时可能会选择发货超过 100 件以保证更多客户能够收到至少 100 件正品。如果你是生产商。为保证 99% 的好评率，请问你应该至少选择发货多少件?

  Solution: 如果发货 100+k 件, 则其中次品的数目近似服从 $Pois((100+k)\times 2.28\%)$。试验得到

$$P\left(Pois(106\times 2.28\%)\leq 6\right)=0.9880$$

以及

$$P\left(Pois(107 \times 2.28\%) \le 7\right) = 0.9963$$

因此至少应该选择发货 107 件。

注：第三问有直接使用二项分布的同学、或者因为第一问次品率近似位数不同而出现估计件数的上下出入，均视为正确答案

(6) 假设你手中只有一枚均匀的硬币

- (10 pt) 试仅通过这枚硬币近似生成一个 $\lambda = 1$ 的泊松分布的随机变量。

  Solution: 对于一个较大的 n，每一组投掷该硬币 n 次，投掷 2^n 组试验。对于任意 $k \le 2^n$，定义随机变量 $X_k = 1$ 当且仅当第 k 组试验中结果全部都是正面。现在令 $S_n = X_1 + X_2 + \cdots + X_{2^n}$ 我们有 S_n 服从二项分布 $B(2^n, 2^{-n})$，当 n 较大时近似于 $\lambda = 1$ 的泊松分布。

- (5 pt) 试仅通过这枚硬币近似生成一个 $\lambda = 1$ 的指数分布的随机变量。（提示：可尝试去寻找概率分布函数（CDF）的近似）、

  Solution: 对于一个较大的 n，每一组投掷该硬币 n 次，投掷 2^n 组试验。对于任意 $k \le 2^n$，定义随机变量 $X_k = 1$ 当且仅当第 k 组试验中结果全部都是正面。现在令 $\bar{T}_n$ 为第一次使得 $X_k = 1$ 的 k，以及 $T_n = \bar{T}_n / 2^n$。我们有

  $$P(T_n \le t) = 1 - P(T_n > t) = 1 - P(\bar{T}_n > \lfloor 2^n t \rfloor) = 1 - (1 - 2^{-n})^{\lfloor 2^n t \rfloor} \approx 1 - e^{-t}$$

  因此 T_n 即满足我们的要求。