

# MATH 505a Spring 2020 Qual Solution Attempts

Troy Tao

July 17, 2022

Contact [yntao@usc.edu](mailto:yntao@usc.edu) if you think this document needs revision.

## Problem 1

Each pack of bubble gum contains one of  $n$  types of coupon, equally likely to be each of the types, independently from one pack to another. Let  $T_j$  be number of packs you must buy to obtain coupons of  $j$  different types. Note that  $T_1 = 1$  always.

(a)

Find the distribution and expected value of  $T_2 - T_1$  and of  $T_3 - T_2$ .

*Solution.* Notice that these differences are geometric distributions:

$$\mathbb{P}(T_2 - T_1 = t) = \left(\frac{n-1}{n}\right) \left(\frac{1}{n}\right)^{t-1}, \quad t \geq 1.$$

$$\mathbb{P}(T_3 - T_2 = t) = \left(\frac{n-2}{n}\right) \left(\frac{2}{n}\right)^{t-1}, \quad t \geq 1.$$

$$\mathbb{E}(T_2 - T_1) = \frac{n}{n-1}$$

$$\mathbb{E}(T_3 - T_2) = \frac{n}{n-2}$$

(b)

Compute  $\mathbb{E}T_n$ .

*Solution.*

$$\begin{aligned} \mathbb{E}(T_n) &= \mathbb{E}[(T_n - T_{n-1}) + (T_{n-1} - T_{n-2}) + \cdots + (T_2 - T_1) + T_1] \\ &= n + \frac{n}{2} + \cdots + \frac{n}{n-1} + 1 \\ &= \sum_{k=1}^n \frac{n}{k} \end{aligned}$$

(c)

Fix  $k$  and let  $A_i$  be the event that none of the first  $k$  packs you buy contain coupon  $i$ . Find  $\mathbb{P}(A_1 \cup A_2 \cup A_3 \cup A_4)$ . Then fix  $\alpha > 0$ , take  $k = \lfloor \alpha n \rfloor$  and find the limit of this probability as  $n \rightarrow \infty$ . Here  $\lfloor x \rfloor$  denotes the largest integer  $\leq x$ . HINT: Consider probabilities  $\mathbb{P}(A_i)$ ,  $\mathbb{P}(A_i \cap A_j)$ , etc.

*Solution.* By inclusion-exclusion theorem,

$$\begin{aligned}\mathbb{P}(A_1 \cup A_2 \cup A_3 \cup A_4) &= \sum_{i=1}^4 \mathbb{P}(A_i) - \sum_{i,j=1}^4 \mathbb{P}(A_i \cap A_j) + \sum_{i,j,k=1}^4 \mathbb{P}(A_i \cap A_j \cap A_k) - \mathbb{P}(A_1 \cap A_2 \cap A_3 \cap A_4) \\ &= \binom{4}{1} \left(\frac{n-1}{n}\right)^k - \binom{4}{2} \left(\frac{n-2}{n}\right)^k + \binom{4}{3} \left(\frac{n-3}{n}\right)^k - \binom{4}{4} \left(\frac{n-4}{n}\right)^k \\ &= 4 \left(1 - \frac{1}{n}\right)^k - 6 \left(1 - \frac{2}{n}\right)^k + 4 \left(1 - \frac{3}{n}\right)^k - \left(1 - \frac{4}{n}\right)^k\end{aligned}$$

Now, take  $k = \lfloor \alpha n \rfloor$  and  $n \rightarrow \infty$ , we have:

$$\begin{aligned}\lim_{n \rightarrow \infty} \mathbb{P}(A_1 \cup A_2 \cup A_3 \cup A_4) &= \lim_{n \rightarrow \infty} \left( 4 \left(1 - \frac{1}{n}\right)^{\lfloor \alpha n \rfloor} - 6 \left(1 - \frac{2}{n}\right)^{\lfloor \alpha n \rfloor} + 4 \left(1 - \frac{3}{n}\right)^{\lfloor \alpha n \rfloor} - \left(1 - \frac{4}{n}\right)^{\lfloor \alpha n \rfloor} \right) \\ &= \lim_{n \rightarrow \infty} \left( 4 \left(1 - \frac{1}{n}\right)^{\alpha n} - 6 \left(1 - \frac{2}{n}\right)^{\alpha n} + 4 \left(1 - \frac{3}{n}\right)^{\alpha n} - \left(1 - \frac{4}{n}\right)^{\alpha n} \right) \\ &= 4e^{-\alpha} - 6e^{-2\alpha} + 4e^{-3\alpha} - e^{-4\alpha}\end{aligned}$$

(d)

Assume there are  $n = 4$  coupon types; find  $\mathbb{P}(T_4 > k)$  for all  $k \geq 4$ . HINT: This is short if you use what you've already done.

*Solution.* From (c), we have:

$$\mathbb{P}(T_4 > k) = 4 \left(\frac{3}{4}\right)^k - 6 \left(\frac{1}{2}\right)^k + 4 \left(\frac{1}{4}\right)^k, \quad k \geq 4$$

## Problem 2

Let  $X$  be exponential( $\lambda$ ) (that is, density  $f(x) = \lambda e^{-\lambda x}$ ). The *integer part* of  $X$  is  $\lfloor X \rfloor = \max\{k \in \mathbb{N} : k \leq X\}$ . The *fractional part* of  $X$  is  $X - \lfloor X \rfloor$ . Show that  $\lfloor X \rfloor$  and  $X - \lfloor X \rfloor$  are independent. HINT: In general, two random variables  $U, V$  are independent if the distribution of  $V$  conditioned on  $U = u$  doesn't depend on  $u$ .

*Proof.* It suffices to show the conditional probability is same as the unconditioned one:

$$\begin{aligned}
\mathbb{P}(\lfloor X \rfloor = n | X - \lfloor X \rfloor = \alpha) &= \frac{f_X(n + \alpha)}{\sum_{k=0}^{\infty} f_X(k + \alpha)} \\
&= \frac{\lambda e^{-\lambda(n+\alpha)}}{\sum_{k=0}^{\infty} \lambda e^{-\lambda(k+\alpha)}} \\
&= \frac{e^{-\lambda n}}{\sum_{k=0}^{\infty} e^{-\lambda k}} \\
&= \frac{e^{-\lambda n}}{\frac{1}{1-e^{-\lambda}}} \\
&= e^{-\lambda n} - e^{-\lambda(n+1)} \\
&\stackrel{(*)}{=} \mathbb{P}(n \leq X < n+1) \\
&= \mathbb{P}(\lfloor X \rfloor = n)
\end{aligned}$$

(\*)  $X$  is continuous. □

### Problem 3

Let  $X_1, X_2, X_3$  be i.i.d. uniform in  $[0,1]$ . Let  $X_{(1)}$  be the smallest of the 3 values,  $X_{(2)}$  the second smallest, and  $X_{(3)}$  the largest.

(a)

Find the distribution function and expected value for  $X_{(1)}$ .

*Solution.* Let  $F_{(i)}$  denotes the cdf of  $X_{(i)}$ .

$$\begin{aligned}
F_{(1)}(x) &= \mathbb{P}(X_{(1)} \leq x) \\
&= 1 - \mathbb{P}(X_{(1)} > x) \\
&= 1 - \prod_{i=1}^3 \mathbb{P}(X_i > x) \\
&= 1 - (1-x)^3, \quad 0 < x < 1
\end{aligned}$$

Since  $X_{(1)} \geq 0$ , we can compute expectation using complementary cdf:

$$\begin{aligned}
\mathbb{E}(X_{(1)}) &= \int_0^1 (1-x)^3 dx \\
&= \frac{1}{4}
\end{aligned}$$

**(b)**

Find the distribution function and the density of  $X_{(2)}$ .

*Solution.*

$$\begin{aligned}\mathbb{P}(X_{(2)} \leq x) &\stackrel{(*)}{=} (\mathbb{P}(X_1 \leq x))^3 + \binom{3}{2} (\mathbb{P}(X_1 \leq x))^2 (\mathbb{P}(X_1 > x)) \\ &= x^3 + 3x^2(1-x), \quad 0 < x < 1 \\ f_{(2)}(x) &= 6x - 6x^2, \quad 0 < x < 1\end{aligned}$$