

MATH 505a Fall 2016 Qual Solution Attempts

Troy Tao

August 4, 2022

Contact yntao@usc.edu if you think this document needs revision.

Problem 1

You have a choice to roll a fair die either 100 times. For each of the following outcomes, state whether it is more likely with 100 rolls, or with 1000 rolls. Justify your answer, but you do not need to give a full formal proof.

(a)

The number 1 shows on the die between 15% and 20% of the time.

Solution. $n = 1000$. (not a proof) Let X_i be the indicator function for the i th roll being 1. By the Weak Law of Large Numbers (WLLN),

$$\frac{1}{n} \sum_{i=1}^n X_i \rightarrow \mathbb{E}(X_i) = \frac{1}{6} \text{ in probability.}$$

That is, for any $\epsilon > 0$,

$$\lim_{n \rightarrow \infty} \mathbb{P}\left(\frac{1}{6} - \epsilon < \frac{1}{n} \sum_{i=1}^n X_i < \frac{1}{6} + \epsilon\right) = 1$$

Notice that $15\% < \frac{1}{6} < 20\%$. Therefore,

$$\lim_{n \rightarrow \infty} \mathbb{P}(15\% < \frac{1}{n} \sum_{i=1}^n X_i < 20\%) = 1$$

In another word, we should expect that the above probability is larger (closer to 1) when n larger.

(b)

The number showing is at most 3, at least half the time.

Solution. $n = 100$.

Proof. Let X_i be the indicator function for the i th roll at most 3. Compute the probability for $2n$ rolls:

$$\begin{aligned} \mathbb{P}\left(\frac{1}{2n}\sum_{i=1}^{2n}X_i \geq \frac{1}{2}\right) &= \sum_{i=n}^{2n} \binom{2n}{i} \left(\frac{1}{2}\right)^{2n} \\ &= \frac{1}{2} \cdot 2 \sum_{i=n}^{2n} \binom{2n}{n} \left(\frac{1}{2}\right)^{2n} \\ &\stackrel{(*)}{=} \frac{1}{2} \left(\binom{2n}{n} \left(\frac{1}{2}\right)^{2n} + \sum_{i=0}^{n-1} \binom{2n}{i} \left(\frac{1}{2}\right)^{2n} \right) \\ &= \frac{1}{2} + \left(\frac{1}{2}\right)^{2n+1} \binom{2n}{n} \end{aligned}$$

(*) By the symmetry that $\binom{n}{m} = \binom{n}{n-m}$.

Compare the above probability between $2n$ and $2n+2$ notice that for the right term,

$$\begin{aligned} \frac{\binom{2n}{n} \left(\frac{1}{2}\right)^{2n+1}}{\binom{2n+2}{n+1} \left(\frac{1}{2}\right)^{2n+3}} &= \frac{4 \cdot (n+1)^2}{(2n+1)(2n+2)} \\ &= \frac{4n^2 + 8n + 4}{4n^2 + 6n + 2} \\ &> 1, \text{ for } n > 0 \end{aligned}$$

Therefore the probability is decreasing with n . □

(c)

The number showing is 2 or 5, at least half the time.

Solution. $n = 100$. (not a proof) Let X_i be the indicator function for i th roll being 2 or 5. By the WLLN,

$$\frac{1}{n} \sum_{i=1}^n X_i \rightarrow \mathbb{E}(X_i) = \frac{1}{3}, \text{ in probability.}$$

Since $\frac{1}{2} > \frac{1}{3}$,

$$\lim_{n \rightarrow \infty} \mathbb{P}\left(\frac{1}{n} \sum_{i=1}^n X_i \geq \frac{1}{2}\right) = 0$$

Therefore, we should expect that this probability is smaller when n is larger.

Problem 2

let X and Y be independent exponential random variables with parameters λ and μ (that is, $\mathbb{E}(X) = 1/\lambda$ and $\mathbb{E}(Y) = 1/\mu$), and let $Z = \min(X, Y)$.

(a)

Show that Z is independent of the event $X < Y$. In other words, show the event $Z \leq t$ is independent of $X < Y$ for all t .

Proof.

$$\begin{aligned}\mathbb{P}(Z \leq t | X < Y) &= \mathbb{P}(X \leq t | X < Y) \\ &= \frac{\mathbb{P}(X \leq t, X < Y)}{\mathbb{P}(X < Y)} \\ \mathbb{P}(X \leq t, X < Y) &= \int \mathbb{P}(x \leq t, x < Y | X = x) f_X(x) dx \\ &\stackrel{(*)}{=} \int \mathbf{1}_{x \leq t} \cdot \mathbb{P}(Y > x) f_X(x) dx \\ &= \int_0^t e^{-\mu x} \cdot \lambda e^{-\lambda x} dx \\ &= \frac{\lambda}{\lambda + \mu} (1 - e^{-(\lambda + \mu)t})\end{aligned}$$

(*) By the independence of X and Y .

$$\begin{aligned}\mathbb{P}(X < Y) &= \int \mathbb{P}(X < y) f_Y(y) dy \\ &= \int_0^\infty (1 - e^{-\lambda y}) \cdot \mu e^{-\mu y} dy \\ &= \frac{\lambda}{\lambda + \mu}\end{aligned}$$

According to what we computed above,

$$\begin{aligned}\mathbb{P}(Z \leq t | X < Y) &= 1 - e^{-(\lambda + \mu)t} \\ \mathbb{P}(Z \leq t) &= 1 - \mathbb{P}(Z > t) \\ &= 1 - \mathbb{P}(Y > t) \mathbb{P}(X > t) \\ &= 1 - e^{-(\lambda + \mu)t}\end{aligned}$$

□

(b)

Find the distribution of $\max(X - Y, 0)$.

Solution. For $t > 0$

$$\begin{aligned}\mathbb{P}(\max(X - Y, 0) \leq t) &= \mathbb{P}(X - Y \leq t, X \geq Y) + \mathbb{P}(0 \leq t, X < Y) \\ &= \int_0^\infty \mathbb{P}(y \leq X \leq y + t) \cdot \mu e^{-\mu y} dy + \mathbb{P}(X < Y) \\ &= \int_0^\infty (e^{-\lambda y} - e^{-\lambda(y+t)}) \mu e^{-\mu y} dy \\ &= \frac{\mu}{\mu + \lambda} (1 - e^{-\lambda t})\end{aligned}$$

Problem 3

Let X_1, X_2, \dots be iid with characteristic function ψ . Let N be independent of the X_i 's with $\mathbb{P}(N = n) = 2^{-n}$ for all $n \geq 1$. Let $Y = \sum_{i=1}^n X_i$. Find the characteristic function of Y .

Solution.

$$\begin{aligned}\mathbb{E}(e^{itY}) &= \mathbb{E}(e^{it \sum_{j=1}^N X_j}) \\ &= \mathbb{E}\left(\prod_{j=1}^N \mathbb{E}(e^{itX_j})\right) \\ &= \mathbb{E}(\psi(t)^N) \\ &= \sum_{n=1}^{\infty} \psi(t)^n 2^{-n} \\ &= \begin{cases} \frac{\psi(t)}{2-\psi(t)}, & |\psi(t)/2| < 1 \\ \infty, & \text{otherwise.} \end{cases}\end{aligned}$$

Problem 4

$n \geq 4$ men, among whom are Alfred, Bill, Charles and David, stand in a row. Assume that all possible orderings of the men are equally likely.

(a)

Find the probability that Charles stands somewhere between Alfred and Bill. (Note this does not mean they are necessarily adjacent—there might be other people between Alfred and Bill.)

Solution. Let A, B, C, D denote Alfred, Bill, Charles, and David, respectively. And let $C_{A,B}$ denotes the event that Charles is between A and B. Now, since all the orderings are equally likely, to find the desired probability, we only need to consider the ordering for A, B , and C . They are: $ABC, ACB, BAC, BCA, CAB, CBA$. So

$$\mathbb{P}(C_{A,B}) = \frac{1}{3}$$

(b)

Find the probability that David stands somewhere between Alfred and Bill given that Charles stands somewhere between Alfred and Bill.

Solution. Again, we only need to consider the relative ordering of A, B, C, D . For the cases that C, D are both between A, B : $ACDB, ADCB, BCDA, BDCA$

$$\begin{aligned}\mathbb{P}(D_{A,B}|C_{A,B}) &= \frac{\mathbb{P}(D_{A,B} \cap C_{A,B})}{\mathbb{P}(C_{A,B})} \\ &= \frac{\frac{4!}{3}}{4!} \\ &= \frac{1}{3}\end{aligned}$$

(c)

Find the expected value and variance of the number of men out of n who stand between Alfred and Bill. (Note Alfred and Bill themselves are not counted in this number.)

Solution. Let X_i be the indicator function of i th person(not including A and B) is between A and B . Use the results from (a) and (b):

$$\begin{aligned}\mathbb{E}\left(\sum_{i=1}^{n-2} X_i\right) &= \sum_{i=1}^{n-2} \mathbb{E}(X_i) \\ &= \sum_{i=1}^{n-2} \mathbb{P}(X_i = 1) \\ &= \frac{n-2}{3}\end{aligned}$$

$$\begin{aligned}\text{Var}\left(\sum_{i=1}^{n-2} X_i\right) &= \mathbb{E}\left(\sum_{i=1}^{n-2} X_i\right)^2 - \left(\mathbb{E}\left(\sum_{i=1}^{n-2} X_i\right)\right)^2 \\ &= \sum_{i=1}^{n-2} \mathbb{E}(X_i^2) + \sum_{i \neq j} \mathbb{E}(X_i X_j) - \left(\frac{n-2}{3}\right)^2 \\ &= \left(\frac{n-2}{3}\right) + \frac{1}{6} \cdot (n-2)(n-3) - \left(\frac{n-2}{3}\right)^2\end{aligned}$$