1. **(11 pt.) Aggregating Guesses**

In this problem, we'll consider several different settings where we are aggregating a large number of noisy, unbiased estimates. Suppose a class has n students. Each student is asked to estimate the current temperature. Assume that they each provide independent, unbiased estimates, with X_i denoting the ith student's guess. Let v_i denote $\text{Var}[X_i]$.

(a) **(4 pt.)** Suppose we know each of the v_i's and decide to compute a weighted combination $Z = \sum_i w_i X_i$, where the weights $w_i \geq 0$ are chosen so as to minimize the variance of Z, subject to $\sum_i w_i = 1$. What are those optimal weights as a function of the v_i's, and roughly how accurate will Z be? Please give an answer of the form: "with probability at least 0.9, Z will be within *blah* of the true temperature, where *blah* is a function of the v_i's.

(b) **(5 pt.)** For this part, assume each X_i is drawn from a normal (Gaussian) distribution, whose mean is the true temperature, and whose variance is 1. Roughly how accurate should we expect the *median* of the n guesses to be? As above, please give an answer of the form: "with probability at least 0.9, the median of the X_i's will be within *blah* of the true temperature," where *blah* is a function of n. Your value of *blah* should be accurate up to a constant factor and use big-Oh notation, for example $O(1/n^{3/4})$ or something like that.

[**HINT:** *The following basic fact about a Gaussian should be helpful, and is the only property of a Gaussian that you will need: if Y is a Gaussian with mean μ and variance 1, for any $\epsilon \in (0, 1/2)$ $\Pr[Y < \mu - \epsilon] = \Pr[Y > \mu + \epsilon] < 1/2 - 0.3\epsilon$.*]

(c) **(2 pt.)** Answer the same question as above for the *mean* of the n values X_i. How do your answers compare?

(d) **(0 pt.)** [**Optional: This is a research-level problem.**] As above, suppose each X_i is independently drawn from a normal distribution whose mean is the true temperature, and variance v_i. Assume you know the (multi)set of the v_i's, but you don't know which variance corresponds to which guess. How well should you expect to do, and is there an efficient algorithm that achieves this?

(e) **(0 pt.)** [**Optional: This is a research-level problem.**] Suppose we are in the setting above, but don't know anything about the variances. What is a near-optimal algorithm, and how well will it do, as a function of the (unknown) list of variances $v_1, \ldots$?

[**HINT:** *Note that if two X_i's are identical (or super, super close) then we know that two of the variances are 0 (or really, really small), and hence either of those X_i's would give an extremely accurate guess, no matter what the other $n-2$ guesses are...*]

SOLUTION:

a) We have $Z = \sum_i w_i X_i$ and $v_i = \text{Var}(X_i)$, so $\text{Var}(Z) = \sum_i w_i^2 v_i$. We impose the constraint $\sum_i w_i = 1$ so that $\mathbb{E}[Z]$ has the correct value. To pick weights that minimize

Var(Z) subject to this constraint, we use Lagrange multipliers. Let

$$\begin{aligned} L(w_1, \ldots, w_n, \lambda) &= \sum_i w_i^2 v_i - \lambda \left(\sum_i w_i - 1 \right) \\ &= \sum_i (w_i^2 v_i - \lambda w_i) + \lambda. \end{aligned}$$

When $0 = \frac{\partial L}{\partial w_i} = 2w_i v_i - \lambda$, we get $w_i = \lambda/(2v_i)$. Using the constraint $\sum_i w_i = 1$, we get $w_i = 1/\left(v_i \sum_j (1/v_j)\right)$. With this choice of weights,

$$\mathrm{Var}(Z) = \sum_i \left(\frac{1}{v_i \sum_j \frac{1}{v_j}} \right)^2 v_i = \left(\frac{1}{\sum_j \frac{1}{v_j}} \right)^2 \left(\sum_i \frac{1}{v_i} \right) = \frac{1}{\sum_i \frac{1}{v_i}}.$$

To provide a concrete bound, we want to find c such that

$$\mathbb{P}(|Z - \mathbb{E}[Z]| \leq c) \geq 0.9.$$

Equivalently, we want

$$\mathbb{P}(|Z - \mathbb{E}[Z]| \geq c) \leq \frac{\mathrm{Var}(Z)}{c^2} = 1 - 0.9.$$

Taking $c = \sqrt{\mathrm{Var}(Z)/0.1}$, we get

$$\mathbb{P}\left(|Z - \mathbb{E}[Z]| \leq \sqrt{\frac{10}{\sum_i \frac{1}{v_i}}} \right) \geq 0.9.$$

b) Fix $0 < \varepsilon < 1/2$. Let Y_i be an indicator random variable that is 1 when $X_i < \mu - \varepsilon$ and 0 otherwise, where $\mu = \mathbb{E}[X_i]$ is the true temperature. Then

$$\mathbb{E}[Y_i] = \mathbb{P}(X_i < \mu - \varepsilon) \leq 1/2 - 0.3\varepsilon.$$

Write $Y = \sum_i Y_i$. The median of the X_i is less than $\mu - \varepsilon$ if and only if $Y > n/2$. At this point we can use Chebyshev's inequality to bound the probability that $Y > n/2$. We have

$$\mathrm{Var}(Y) = \sum_i \mathrm{Var}[Y_i] = np(1-p) \leq n,$$

where $p \leq 1/2 - 0.3\varepsilon$ is the probability that Y_i is equal to 1. Thus, by Chebyshev's inequality,

$$\begin{aligned} \mathbb{P}[Y > n/2] &\leq \mathbb{P}[(Y - \mathbb{E}Y) > 0.3\epsilon n \\ &\leq \frac{\mathrm{Var}[Y]}{(0.3\epsilon)^2 n^2} \\ &\leq \frac{1}{(0.3\epsilon)^2 n} \end{aligned}$$

Thus, if we take $\epsilon = c/\sqrt{n}$ for $c = \sqrt{20}/0.3$, we can guarantee that this is at most 0.05. In particular, with probability at least 0.95, the median is no smaller than $\mu - c/\sqrt{n}$. The same argument shows that with probability at least 0.95, the median is no larger than $\mu + c/\sqrt{n}$, and altogether, with probability at least 0.9, the median is within $O(1/\sqrt{n})$ of μ.

Alternate solution: We could also use a Chernoff bound to bound the probability that $Y > n/2$. That would look like this:

$$\begin{aligned}
\mathbb{P}\left(Y > \frac{n}{2}\right) &= \mathbb{P}\left(Y > \frac{n/2}{\mathbb{E}[Y]}\mathbb{E}[Y]\right) \\
&\leq \mathbb{P}\left(Y > \frac{n/2}{n/2 - 0.3\varepsilon n}\mathbb{E}[Y]\right) \\
&= \mathbb{P}\left(Y > \left(1 + \frac{0.3\varepsilon}{1/2 - 0.3\varepsilon}\right)\mathbb{E}[Y]\right) \\
&\leq \exp\left(-\mathbb{E}[Y]\left(\frac{0.3\varepsilon}{1/2 - 0.3\varepsilon}\right)^2 /3\right) && \text{(Lecture 5, Corollary 5)} \\
&= \exp\left(-\frac{n(0.3\varepsilon)^2}{3(1/2 - 0.3\varepsilon)}\right) \\
&= \exp(-O(n\varepsilon^2)).
\end{aligned}$$

Setting $\exp(-O(n\varepsilon^2))$ equal to a constant and solving for ε, there is a 0.9 probability that the median of the X_i is no less than $\mu - O(1/\sqrt{n})$. With the same argument, we get the same bound for the median being no larger than $\mu + O(1/\sqrt{n})$.

c) Let $X = \frac{1}{n}\sum_i X_i$ be the average of the X_i. Then we can compute the variance of X as

$$\text{Var}[X] = \frac{1}{n^2}\sum_i \text{Var}(X_i) = \frac{1}{n}.$$

Therefore, by Chebyshev's inequality,

$$\mathbb{P}[|X - \mu| > \epsilon] \leq \frac{1/n}{\epsilon^2}.$$

In particular, if $\epsilon = \sqrt{10/n}$, this is at most 0.1, and we conclude that with probability at least 0.9, X is within $O(1/\sqrt{n})$ of the true mean μ. So we get similar behavior as we did with the median.

2. **(11 pt.) Concentration without Independence**

A computer system has n different failure modes, each of which happens with a small probability. Fortunately, the system is designed to be sufficiently robust in the following sense: as long as less than half of the failures occur, things are fine; otherwise, a large-scale crash will happen. We want to make sure that the probability of crashing is small enough.

To model the above scenario, we define n Bernoulli random variables $X_1, \ldots, X_n$. Each X_i

is the indicator of the i-th failure mode, i.e., $X_i = 1$ if failure i occurs and $X_i = 0$ otherwise. Our goal is to upper bound the probability $\Pr\left[\sum_{i=1}^n X_i \geq n/2\right]$.

(a) **(2 pt.)** Let's first assume that the n failure events are independent and the probability of each failure is at most $1/3$. Formally, we have:

Assumption 1. $\Pr[X_i = 1] \leq 1/3$ *for every* $i \in [n]$ *and* $X_1, \ldots, X_n$ *are independent.*

Prove that under Assumption 1, for some constant $C > 0$ that does not depend on n,

$$\Pr\left[\sum_{i=1}^n X_i \geq n/2\right] \leq e^{-Cn}. \tag{1}$$

Thus, the probability of a crash is exponentially small in n.

[**HINT:** *Feel free to use (without proof) any of the Chernoff bounds in lecture note #5 (including Theorem 2 and Corollaries 5 and 6) and also the inequality* $\frac{e^\delta}{(1+\delta)^{1+\delta}} \leq e^{-\delta^2/3}$ *for* $\delta \in [0, 1]$.]

(b) **(1 pt.)** Now we decide that Assumption 1 is too unrealistic, since many of the failure modes are known to be strongly correlated. Show that only assuming $\Pr[X_i = 1] \leq 1/3$ (and not the independence), the probability of crashing can be as large as $\Omega(1)$. In particular, prove that for any $n \geq 1$, there exist random variables $X_1, \ldots, X_n$ that satisfy: (1) $\Pr[X_i = 1] \leq 1/3$ for every $i \in [n]$; (2) $\Pr\left[\sum_{i=1}^n X_i \geq n/2\right] \geq 1/3$.

(c) **(2 pt.)** Let's try the following relaxation of Assumption 1, which states that the probability for k different failures to occur simultaneously is exponentially small in k:

Assumption 2. *For any* $S \subseteq [n]$, $\Pr\left[X_i = 1 \text{ for all } i \in S\right] \leq (1/3)^{|S|}$.

Show that Assumption 2 is strictly weaker than Assumption 1 by proving: (1) Assumption 1 implies Assumption 2; (2) the implication on the other direction does not hold, i.e., there exist some n and $X_1, \ldots, X_n$ that satisfy Assumption 2 but not Assumption 1.

[**HINT:** *For (2), there exists a counterexample for* $n = 2$.]

(d) **(6 pt.)** Prove that under Assumption 2, inequality (1) holds for some constant $C > 0$. In your proof, you can appeal to the proof of the Chernoff bounds from lecture videos/notes if you need to write it out verbatim at some point. For example, if you manage to upper bound $\Pr\left[\sum_{i=1}^n X_i \geq n/2\right]$ by an expression involving the moment-generating function of some random variable Y that is the sum of n independent Bernoulli random variables, you can simply say that "the rest of the proof is exactly the proof of Theorem 2 from Lecture #5".

[**HINT:** *Consider independent Bernoulli random variables* $Y_1, \ldots, Y_n$ *with* $Pr[Y_i = 1] = 1/3$ *for each* $i \in [n]$. *For distinct indices* $i, j, \ell \in [n]$, *does* $\mathbb{E}\left[X_i X_j X_\ell\right] \leq \mathbb{E}\left[Y_i Y_j Y_\ell\right]$ *hold? Can you extend your proof of the inequality to the case with repeating indices?*]

[**HINT:** *Let* $X = \sum_{i=1}^n X_i$ *and* $Y = \sum_{i=1}^n Y_i$. *What can we say about* $\mathbb{E}\left[X^k\right]$ *and* $\mathbb{E}\left[Y^k\right]$ *for integer* $k \geq 0$? *Considering the identity* $e^z = \sum_{k=0}^{+\infty} \frac{z^k}{k!}$, *what can we say about* $\mathbb{E}\left[e^{tX}\right]$ *and* $\mathbb{E}\left[e^{tY}\right]$ *for any* $t > 0$?]

(e) **(0 pt.)** **[Optional: this won't be graded.]** Can you construct counterexamples for Part 2b that satisfy *pairwise independence* but have a crashing probability of $\Omega(1/n)$?

Formally, prove that there exists $C > 0$ such that for any $n \geq 2$, there exist $X_1, \ldots, X_n$ that satisfy: (1) $\Pr[X_i = 1] \leq 1/3$; (2) X_i and X_j are independent for distinct $i, j \in [n]$; (3) $\Pr\left[\sum_{i=1}^n X_i \geq n/2\right] \geq C/n$.

[**NOTE:** *This shows that unlike Chebyshev's inequality, Chernoff bounds do not hold if we only assume pairwise independence.*]

[**HINT:** *Recall pairwise independent hash functions if you have seen them before. You can use the Bertrand-Chebyshev theorem, which states that for any integer $n \geq 1$, there exists a prime number p with $n < p < 2n$.*]

SOLUTION:

(a) Let $X = \sum_{i=1}^n X_i$. Since $\mathbb{E}[X_i] = \Pr[X_i = 1] \leq 1/3$ for each $i \in [n]$, $\mathbb{E}[X] = \sum_{i=1}^n \mathbb{E}[X_i] \leq n/3$. Applying Corollary 6 from Lecture #5 with $c = n/3$ and $\delta = 1/2$ gives

$$\Pr\left[\sum_{i=1}^n X_i \geq n/2\right] = \Pr[X \geq (1+\delta)c] \leq \left(\frac{e^\delta}{(1+\delta)^{1+\delta}}\right)^c \leq \exp(-\delta^2 c/3) = \exp(-n/36).$$

The third step applies $\frac{e^\delta}{(1+\delta)^{1+\delta}} \leq e^{-\delta^2/3}$ since $\delta \in [0,1]$. Therefore, (1) holds for $C = 1/36$.

(b) Consider random variables $X_1, \ldots, X_n$ with the following property: with probability $1/3$, they all take value 1; and with the remaining probability $2/3$, they all take value 0. It is easy to check that

$$\Pr[X_1 = 1] = \cdots = \Pr[X_n = 1] = \Pr\left[\sum_{i=1}^n X_i \geq n/2\right] = \Pr[X_1 = \cdots = X_n = 1] = 1/3,$$

as desired.

(c) To show that Assumption 1 implies Assumption 2, we use the independence of X_i's to get $\Pr[X_i = 1 \text{ for all } i \in S] = \prod_{i \in S} \Pr[X_i = 1]$. This product is at most $(1/3)^{|S|}$ because every factor $\Pr[X_i = 1]$ is at most $1/3$.

To show that Assumption 2 does not imply Assumption 1, we consider random variables $X_1, \ldots, X_n$ with the following property: with probability $(1/3)^n$, they all take value 1; and with the remaining probability $1 - (1/3)^n$, they all take value 0. For any non-empty $S \subseteq [n]$, we have $\Pr[X_i = 1 \text{ for all } i \in S] = (1/3)^n \leq (1/3)^{|S|}$, so Assumption 2 is satisfied. However, when $n \geq 2$, $\Pr[X_1 = X_2 = 1] = (1/3)^n \neq (1/3)^{2n} = \Pr[X_1 = 1]\Pr[X_2 = 1]$, so X_1 is not independent from X_2, and thus Assumption 1 is not satisfied.

(d) Following the hint, we define independent Bernoulli random variables $Y_1, \cdots, Y_n$ with $\Pr[Y_i = 1] = 1/3$. Define $X = \sum_{i=1}^n X_i$ and $Y = \sum_{i=1}^n Y_i$. We will prove later that for any $t > 0$, we have

$$\mathbb{E}\left[e^{tX}\right] \leq \mathbb{E}\left[e^{tY}\right]. \tag{2}$$

Assuming that this is true, for every $t > 0$, applying Markov's inequality to the non-negative random variable e^{tX}, we have

$$\Pr[X \geq n/2] = \Pr\left[e^{tX} \geq e^{tn/2}\right] \leq \mathbb{E}\left[e^{tX}\right]/e^{tn/2} \leq \mathbb{E}\left[e^{tY}\right]/e^{tn/2},$$

which implies that $\Pr[X \geq n/2] \leq \inf_{t>0} \mathbb{E}\left[e^{tY}\right]/e^{tn/2}$. Our solution to Part (a), together with the proof of Corollary 6 from Lecture #5, implicitly proved that the infimum on the right-hand side is at most e^{-Cn} for $C = 1/36$. Therefore, we have $\Pr[X \geq n/2] \leq e^{-Cn}$ as desired.

We now turn to proving inequality (2). Because e^{tz} can be expanded as $\sum_{k=0}^{+\infty}(t^k/k!)z^k$ for any $z \in \mathbb{R}$, by the linearity of expectation, we have $\mathbb{E}\left[e^{tX}\right] = \sum_{k=0}^{+\infty}(t^k/k!)\,\mathbb{E}\left[X^k\right]$ and $\mathbb{E}\left[e^{tY}\right] = \sum_{k=0}^{+\infty}(t^k/k!)\,\mathbb{E}\left[Y^k\right]$. It thus suffices to prove that for every $k = 0, 1, \cdots$, we have $\mathbb{E}\left[X^k\right] \leq \mathbb{E}\left[Y^k\right]$. Using the definition $X = \sum_{i=1}^{n} X_i$, we have $X^k = \sum_\sigma \prod_{j=1}^{k} X_{\sigma(j)}$, where the summation is over all n^k functions $\sigma : [k] \to [n]$. By the linearity of expectation again, we have $\mathbb{E}\left[X^k\right] = \sum_\sigma \mathbb{E}\left[\prod_{j=1}^{k} X_{\sigma(j)}\right]$. Similarly, we have $\mathbb{E}\left[Y^k\right] = \sum_\sigma \mathbb{E}\left[\prod_{j=1}^{k} Y_{\sigma(j)}\right]$. It now suffices to prove that for every σ, $\mathbb{E}\left[\prod_{j=1}^{k} X_{\sigma(j)}\right] \leq \mathbb{E}\left[\prod_{j=1}^{k} Y_{\sigma(j)}\right]$. Define $S = \{i \in [n] : \exists j \in [k], \sigma(j) = i\}$ as the image of σ. We have

$$\begin{aligned}
\mathbb{E}\left[\prod_{j=1}^{k} X_{\sigma(j)}\right] &= \Pr[X_i = 1 \text{ for all } i \in S] \\
&\leq (1/3)^{|S|} && \text{(by Assumption 2)} \\
&= \Pr[Y_i = 1 \text{ for all } i \in S] \\
&= \mathbb{E}\left[\prod_{j=1}^{k} Y_{\sigma(j)}\right],
\end{aligned}$$

as desired.

(e) Let p be the minimum prime that is greater than n. Define random variables $Z_1, \ldots, Z_p$ as follows: (1) pick $a, b \in \{0, 1, \ldots, p-1\}$ independently and uniformly at random; (2) set $Z_i = (a \cdot i + b) \bmod p$. We can verify that $Z_1, \ldots, Z_p$ defined above are pairwise independent, since for any $i \neq j$ and $c, d \in \{0, 1, \ldots, p-1\}$, it holds that

$$Z_i = c \text{ and } Z_j = d \iff a = (c-d) \cdot (i-j)^{-1} \bmod p \text{ and } b = (c - a \cdot i) \bmod p.$$

Here $(i-j)^{-1}$ exists since p is a prime and $(i-j) \bmod p \neq 0$. Thus, (Z_i, Z_j) is uniformly distributed among $\{0, 1, \ldots, p-1\}^2$ for $i \neq j$, and thus Z_i and Z_j are independent.

Define $X_1, \ldots, X_n$ such that X_i is the indicator of $Z_i < \lfloor p/3 \rfloor$. It then follows that $X_1, \ldots, X_n$ are pairwise independent and $\Pr[X_i = 1] = \Pr[Z_i < \lfloor p/3 \rfloor] \leq 1/3$. On the other hand, note that when we happen to choose $a = 0$ and $b < \lfloor p/3 \rfloor$, we have $Z_1 = \cdots = Z_p = b < \lfloor p/3 \rfloor$ and thus $X_1 + \cdots + X_n = n \geq n/2$. This indicates that

$$\Pr[X_1 + \cdots + X_n \geq n/2] \geq \Pr[a = 0] \cdot \Pr[b < \lfloor p/3 \rfloor] = \frac{1}{p} \cdot \frac{\lfloor p/3 \rfloor}{p} > \frac{1}{2n} \cdot \frac{1}{5} = \frac{1}{10n}.$$

The third step applies: (1) $1/p > 1/(2n)$, since the Bertrand-Chebyshev theorem guarantees $p < 2n$; (2) $\lfloor p/3 \rfloor \geq p/5$ for every prime number $p \geq 3$, which can be easily verified. Therefore, the lower bound holds with $C = 1/10$.

3. **(8 pt.) Processes and CPUs**

Suppose that in a distributed system, we have N CPUs and P processes. Each process is independently and uniformly allocated to a CPU. However, if multiple processes are allocated to the same CPU, the CPU will choose one of them at random to complete; the remaining processes allocated to that CPU will not be completed.

(a) **(2 pt.)** What is expected number of processes that will be completed?

(b) **(6 pt.)** Suppose that $P \geq N$, and denote the total number of completed processes by C. **Use Poissonization** to prove that $\Pr\left[C \leq \frac{N}{2}\right] \leq e^{-\Omega(N)}$.

[**Note:** There may be ways to do this problem that don't involve Poissonization, but we want you to use it to get practice with it. That is, you should prove the statement by analyzing the case when the number of processes is an appropriate Poisson random variable. Don't forget the de-Poissonization step!]

(c) **(0 pt.)** **[Optional: this won't be graded.]** Let μ be your answer from part (a). Under what conditions on P and N can you use Poissonization to show that $\Pr[C \leq (1-\delta)\mu] \leq e^{-\Omega_\delta(N)}$ for any $\delta > 0$, where the Ω_δ notation means that you are allowed to have constants that depend on δ hidden inside the big-Omega.

SOLUTION:

a) Let C_i be the number of completed processes at CPU i, which is just 1 if CPU i has any processes allocated to it and 0 otherwise. We have $\Pr[C_i = 0] = (\frac{N-1}{N})^P$. Therefore by linearity of expectation $\mathbb{E}[C] = N(1-(\frac{N-1}{N})^P)$. How should you think of this quantity? If $P = N$, then $((N-1)/N)^P = (1-1/N)^N \approx 1/e \approx 0.37$, in which case $\mathbb{E}[C] \approx 0.63N$.. If $P > N$ the expectation is larger than this, and if $P < N$ the expectation is smaller, but will be at least $0.63P$ since the function α^P is convex for $\alpha \in (0,1)$.

b) The high-level intuition for this part is as follows: If, instead of exactly P processes, the number of processes is drawn from a Poisson distribution, then the number of processes allcoated to each of the N CPUs is independent. If The expectation of the Poisson process is close to N, then the expected number of CPUs that have at least one process assigned to it will be roughly the same as in the case where $P = N$, which is $> 0.6N$. Because of independence in the Poisson case, we can apply a Chernoff bound, which will give a probability that we end up with less than $N/2$ nonempty CPUs to be inverse exponential in N, since we are talking about an expectation of at least $\Theta(N)$ from an expectation of $\Theta(N)$. To "depoissonize", we just need to note that the probability we have fewer than $N/2$ CPUs with processes is monotonically decreasing as the number of processes increases (adding extra processes can never decrease the number of non-empty CPUs). So, if we pick the expectation of the Poisson distribution such that the probability it is less than N is, say, at least $1/2$, then the probability of getting less than $N/2$ processes can be at most $1/(1/2) = 2$ times larger in the case with exactly N processes, versus the Poissonized setting. Below we make this a bit more formal.

If, instead of exactly P processes, the number of processes is drawn according to a Poisson distribution of expectation Q, then the number of processes allocated to each of the N CPUs is independent, each distributed according to a Poisson random variable with expectation Q/N. In this Poissonized setting, the event that the ith CPU has no assigned processes is $e^{-Q/N}$, namely the probability this Poisson random variable is equal to 0, and this event is independent of whatever happens with the other CPUs. Hence, in this Poissonized setting, the number of completed processes can be modeled as a sum of N i.i.d. random coin tosses, each of which occurs with probability $1 - e^{-Q/N}$..

Lets let $Q = 0.9N$, and hence the number of completed processes in this Poissonized setting, X, corresponds to a sum of N i.i.d. coin flips, each of which is heads with probability at least $1-e^{-0.9} > 0.59$.. Since $E[X] > 0.59N$, the probability $X \leq N/2$ is at most $Pr[X \leq (1 - 1/6)E[X]] \leq e^{-E[X](1/6)^2/2} = e^{-\Omega(N)}$, where we used the Chernoff bound of Corollary 5.

Now, all that remains is to related this probability in the "Poissonized" setting back to the original setting. To do this, note that the number of completed processes can only increase as the number of processes increases. Hence the probability of failure given N processes is at most the probability of failure in the Poissonized setting, given that the number of processes is at most N, which is at most $e^{-\Omega(N)}/Pr[Z \leq N]$, where Z is a Poisson random variable of expectation N. $Pr[Z \leq N]$ is actually 1 minus an inverse exponential in N, but its good enough to bound this by 1/2 [which is trivially true since the variance is $0.9N$, and hence we get a bound of $O(1/N^2) \ll 1/2$ just from Chebyshev's inequality].