I worked alone without any help.

1 Random Cuckoo Hashing

(a) $\mathbb{P}[\text{No Collision}] = \frac{n!}{n^n}; \longrightarrow 0$

Since the size of the event space of such a situation is $|\omega| = n!$, and the size of the probability space is $|\Omega| = n^n$, so the probability of such a situation is $\mathbb{P}[\text{No Collision}] = \frac{|\omega|}{|\Omega|} = \frac{n!}{n^n}$, and as $n \to \infty$, we can see that $\mathbb{P}[\text{No Collision}] \to 0$.

We can represent $\mathbb{P}[\text{No Collision}]$ in another method: $\mathbb{P}[\text{No Collision}] = \frac{n}{n}\frac{n-1}{n}\cdots\frac{1}{n}$, which will tend toward 0 (all terms of this product are smaller than or equal to 1, and $\lim_{n\to\infty}\frac{1}{n} = 0$) as n grows very large.

(b) $\mathbb{E}[\text{Collisions}] = n - 1$

Let the expected number of collisions that we'll see while hashing D_n be $\mathbb{E}[\text{Collisions for } D_n] = X$.

Since we have already hashed $D_1, \ldots, D_{n-1}$, and they each occupy their own bucket, so in this situation, the probability of D_n not getting a collision is $\frac{1}{n}$ (which is equivalent to having 0 collisions); then, in other words, the probability of D_n getting a first collision is $1 - \frac{1}{n} = \frac{n-1}{n}$.

Now, let D_n take the i^{th} bucket, the bucket of D_i. So now, we reached the same situation where $(n-1)$ pieces of data have occupied their own buckets, and a single piece of data D_i needs to be rehashed, and thus, the expected number of collisions we'll see hashing D_i would be X again, because it's an identical situation. This implies that the total number of collision of hashing D_n in this situation would be $\mathbb{E}[\text{First Collision}] = 1 + X$.

Thus, looking back at $\mathbb{E}[\text{Collisions for } D_n]$, we have this equation:

$$X = \frac{1}{n}\cdot 0 + \frac{n-1}{n}\cdot(1 + X)$$

Thus, we can calculate that:

$$X = n - 1$$

2 Markov's Inequality and Chebyshev's Inequality

(a) True (Direct Proof)

We proceed by a direct proof. Using Theorem 16.1, we have that $\text{var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$, which gives us that:

$$\mathbb{E}[X^2] = \text{var(X)} + \mathbb{E}[X]^2 = 9 + 2^2 = 13$$

This gives the value we desire. Q.E.D.

(b) True (Proof by Contradiction)

Assume, for a contradiction, that $\mathbb{P}[X \leq 1] > 8/9$. Let R denote the assertion that $\mathbb{E}[X] = 2$. Since the value of X is never greater than 10, so we have that

$$\mathbb{E}[X] < 1 \cdot 8/9 + 10 \cdot (1 - 8/9) = 2$$

In other words, the expectation of X would never reach 2, which implies $\neg R$. So $R \wedge \neg R$ holds, which gives the contradiction. Q.E.D.

(c) True (Direct Proof)

Givent that $\mu = \mathbb{E}[X] = 2$ and $\text{var}(X) = 9$, so using Theorem 18.3 (Chebyshev's Inequality), we have that:

$$\mathbb{P}[|X - \mu| \geq 4] \ \leq \ \frac{\text{var}(X)}{4^2}$$

which gives that $\mathbb{P}[|X - 2| \geq 4] \leq \frac{9}{16}$. Then, since $\mathbb{P}[|X - 2| \geq 4] = \mathbb{P}[X \geq 6] + \mathbb{P}[X \leq -2]$, and since $\mathbb{P}[X \leq -2]$ is nonnegative, so we can conclude that:

$$\mathbb{P}[X \geq 6] \ \leq \ \mathbb{P}[|X - 2| \geq 4] \ \leq \ \frac{9}{16}$$

This gives the result we desire. Q.E.D.

(d) False (Counterexample)

Consider random variable X where $\mathbb{P}[X = 6] = \frac{11}{32}, \mathbb{P}[X = -10] = \frac{1}{160}, \mathbb{P}[X = 0] = \frac{13}{20}$.

Here, we can first verify that all values X can take on is not greater than 10. Then, $\mathbb{E}[X] = 6 \cdot \frac{11}{32} + (-10) \cdot \frac{1}{160} + 0 \cdot \frac{13}{20} = 2$ as desired, and $\mathbb{E}[X^2] = 6^2 \cdot \frac{11}{32} + (-10)^2 \cdot \frac{1}{160} + 0^2 \cdot \frac{13}{20} = 13$, which gives that $\text{var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = 13 - 2^2 = 9$ as desired. In other words, this is a random variable that satisfies all constraints. Yet,

$$\mathbb{E}[X \geq 6] \geq \mathbb{E}[X = 6] = \frac{11}{32} > \frac{9}{32}$$

Thus, this is a valid counterexample. Q.E.D.

3 Easy A's

$\mathbb{E} = 35; \text{var} = 25$

Since each of the first 3 problems (and the last 4 problems) are graded independently with the same range, expectation and variance, so the two sets are i.i.d., which implies that I could calculate desired results separately. Let the grading of one of the first three problems be I_1 and let the grading of one of the last four problems be I_2, and denote my total score as the random variable X, where $X = 3I_1 + 4I_2$. As analyzed earlier, we have that the I_i's are mutually independent:

Using given information, we have that $\mathbb{E}[I_i] = 5$ and $\text{var}(I_i) = 1$. Thus, using results from the Notes (Theorem 15.1 and 16.3), we have that:

$$\mathbb{E}[X] = \mathbb{E}[3I_1 + 4I_2] = 3\mathbb{E}[I_1] + 4\mathbb{E}[I_2] = 3 \cdot 5 + 4 \cdot 5 = 35$$

$$\text{var}[X] = \text{var}[3I_1 + 4I_2] = \text{var}[3I_1] + \text{var}[4I_2] = 3^2\text{var}[I_1] + 4^2\text{var}[I_2] = 9 \cdot 1 + 16 \cdot 1 = 25$$

Now, using Chebyshev's inequality, we can calculate that:

$$\mathbb{P}\big[|X - \mathbb{E}[X]| \geq 25\big] \;\leq\; \frac{\text{var}(X)}{25^2}$$

$$\implies \mathbb{P}\big[|X - 35| \geq 25\big] \;\leq\; \frac{25}{625} = \frac{1}{25} < 5\%$$

In other words, I have less than a 5% chance of getting an A when the grades are randomly chosen this way. Q.E.D.

4 Confidence Interval Introduction

(a) $\frac{\sigma^2}{\epsilon^2}$

Since $\sigma = \sqrt{\text{var}(X)}$, so $\text{var}(X) = \sigma^2$. Thus, using Chebyshev's Inequality, we have that the upper bound would be:

$$\mathbb{P}[|X-\mu| \geq \epsilon] \;\leq\; \frac{\text{var}(X)}{\epsilon^2} = \frac{\sigma^2}{\epsilon^2}$$

(b) Direct Proof

Since the event $|X-\mu| < \epsilon$ is equivalent to the event $-\epsilon < X-\mu < \epsilon$, which is then equivalent to the event $X-\epsilon < \mu < X+\epsilon$ by properties of inequalities, which is then equivalent to the event $\mu \in (X-\epsilon, X+\epsilon)$ by definition, so the event space of $|X-\mu| < \epsilon$ is the same as the event space of $\mu \in (X-\epsilon, X+\epsilon)$, and they have the same sample space.

Thus, $\mathbb{P}[|X-\mu| < \epsilon] \;=\; \mathbb{P}\big\{\mu \in (X-\epsilon, X+\epsilon)\big\}$, as desired.

Q.E.D.

(c) $\epsilon = 2\sqrt{5}\sigma$

We wish to have $\mathbb{P}\big\{\mu \in (X-\epsilon, X+\epsilon)\big\} \geq 95\%$. Using our results from parts (a) and (b), so we have that $\mathbb{P}\big\{\mu \in (X-\epsilon, X+\epsilon)\big\} = \mathbb{P}[|X-\mu| < \epsilon] = \mathbb{P}[\overline{|X-\mu| \geq \epsilon}] = 1 - \mathbb{P}[|X-\mu| \geq \epsilon] \geq 1 - \frac{\sigma^2}{\epsilon^2}$. This means that if we can choose ϵ such that $1 - \frac{\sigma^2}{\epsilon^2} \geq 95\%$, then we can guarantee that

$$\mathbb{P}\big\{\mu \in (X-\epsilon, X+\epsilon)\big\} \;\geq\; 1 - \frac{\sigma^2}{\epsilon^2} \;\geq\; 95\%$$

Thus, we can calculate that for $1 - \frac{\sigma^2}{\epsilon^2} \geq 95\%$, so we need:

$$\epsilon^2 \geq 20\sigma^2$$

$$\implies \epsilon \geq 2\sqrt{5}\sigma$$

(d) $\mathbb{E}[\overline{X}] = \mu$, $\text{var}(\overline{X}) = \frac{\sigma^2}{n}$

Since we're given that $n \in \mathbb{Z}^+$ is a constant, and that μ is the mean for X (i.e. $\mu = \mathbb{E}[X]$), and that $X_1, \ldots, X_n$ are i.i.d. samples, as well as that $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$, so we can utilize Theorem 15.1 to get that:

$$\mathbb{E}[\overline{X}] = \mathbb{E}[\frac{1}{n}\sum_{i=1}^{n} X_i] = \frac{1}{n} \cdot \mathbb{E}[X_1 + \cdots + X_n] = \frac{1}{n} \cdot (\mathbb{E}[X_1] + \cdots + \mathbb{E}[X_n]) = \frac{1}{n} \cdot (\mu + \cdots + \mu)$$

$$\implies \mathbb{E}[\overline{X}] = \frac{1}{n} \cdot (n\mu) = \mu$$

And also, using Theorem 16.3 and a result from Note 16 we have:

$$\text{var}(\overline{X}) = \text{var}(\frac{1}{n}\sum_{i=1}^{n} X_i) = (\frac{1}{n})^2 \cdot \text{var}(X_1 + \cdots + X_n) = \frac{1}{n^2} \cdot \big(\text{var}(X_1) + \cdots + \text{var}(X_n)\big)$$

Then, since $\text{var}(X) = \sigma^2$ by definition, so:

$$\text{var}(\overline{X}) = \frac{1}{n^2} \cdot (\sigma^2 + \cdots + \sigma^2) = \frac{1}{n^2} \cdot (n\sigma^2) = \frac{\sigma^2}{n}$$

(e) $\epsilon = \sqrt{\frac{20\sigma^2}{n}}$

We can repeat the process of parts (a) to (c) to choose a proper width ϵ of the confidence interval.

First, denoting the mean of $\overline{X}$ as $\mathbb{E}[\overline{X}] = \nu$, and we calculate an upper bound on $\mathbb{P}\big[|\overline{X} - \nu| \geq \epsilon\big]$, which, using Chebyshev's Inequality, is:

$$\mathbb{P}\big[|\overline{X} - \nu| \geq \epsilon\big] \;\leq\; \frac{\text{var}(\overline{X})}{\epsilon^2} = \frac{\frac{\sigma^2}{n}}{\epsilon^2} = \frac{\sigma^2}{n \cdot \epsilon^2}$$

Now, since the event $|\overline{X} - \nu| < \epsilon$ is equivalent to the event $-\epsilon < \overline{X} - \nu < \epsilon$, which is then equivalent to the event $\overline{X} - \epsilon < \nu < \overline{X} + \epsilon$ by properties of inequalities, which is then equivalent to the event of $\nu \in (\overline{X} - \epsilon, \overline{X} + \epsilon)$ by definition, so the event space of $|\overline{X} - \nu| < \epsilon$ is the same as the event space of $\nu \in (\overline{X} - \epsilon, \overline{X} + \epsilon)$. Thus, $\mathbb{P}[|\overline{X} - \nu| < \epsilon] \;=\; \mathbb{P}\big\{\nu \in (\overline{X} - \epsilon, \overline{X} + \epsilon)\big\}$.

Then, $\mathbb{P}\big\{\nu \in (\overline{X} - \epsilon, \overline{X} + \epsilon)\big\} \;=\; \mathbb{P}[|\overline{X} - \nu| < \epsilon] = \mathbb{P}[\overline{|\overline{X} - \nu| \geq \epsilon}] = 1 - \mathbb{P}[|\overline{X} - \nu| \geq \epsilon]$. Since $\mathbb{P}\big[|\overline{X} - \nu| \geq \epsilon\big] \leq \frac{\sigma^2}{n \cdot \epsilon^2}$, so $\mathbb{P}\big\{\nu \in (\overline{X} - \epsilon, \overline{X} + \epsilon)\big\} = 1 - \mathbb{P}[|\overline{X} - \nu| \geq \epsilon] \;\geq\; 1 - \frac{\sigma^2}{n \cdot \epsilon^2}$.

Since we wish to have $\mathbb{P}\big\{\nu \in (\overline{X} - \epsilon, \overline{X} + \epsilon)\big\} \geq 95\%$, so if $1 - \frac{\sigma^2}{n \cdot \epsilon^2} \geq 95\%$, then we can guarantee our desired result. With σ being known, so we can calculate:

$$1 - \frac{\sigma^2}{n \cdot \epsilon^2} \geq 95\%$$

$$\Longrightarrow \frac{\sigma^2}{n \cdot \epsilon^2} \leq 0.05$$

$$\Longrightarrow 0.05\epsilon^2 \geq \frac{\sigma^2}{n}$$

$$\Longrightarrow \epsilon^2 \geq \frac{20\sigma^2}{n}$$

$$\Longrightarrow \epsilon \geq \sqrt{\frac{20\sigma^2}{n}}$$

Thus, $\epsilon = \sqrt{\frac{20\sigma^2}{n}}$ is an appropriate width of the confidence interval for the desired result, i.e. guaranteeing $\mathbb{P}\big\{\nu \in (\overline{X} - \epsilon, \overline{X} + \epsilon)\big\} \geq 95\%$.

(Confirmed: as n increases, ϵ decreases.)