A.A. The expectation $E(T)$ is our average total cost. We might set our baseline budget to match $E(T)$. The variance $V(T)$ is a measure of how our total cost fluctuates about that average over the years. It is important in managing risk. For example, we might consider the possibility that our total cost could get as large as $E(T) + 2\sqrt{V(T)}$. If T were approximately normal, then budgeting at this level would handle about 97.5% of years.

A.B. Whether or not the X_i are independent, linearity of expectation says that

$$E(T) = E(X_1 + \cdots + X_n) = E(X_1) + \cdots + E(X_n) = \mu_1 + \cdots + \mu_n.$$

A.C. The answer is identical to the answer for A.B.

A.D. If the X_i are independent, then their variances add:

$$V(T) = V(X_1 + \cdots + X_n) = V(X_1) + \cdots V(X_n) = \sigma_1^2 + \cdots + \sigma_n^2.$$

A.E. More generally,

$$V(T) = V(X_1 + \cdots + X_n) = \sum_{i=1}^{n} V(X_i) + 2\sum_{i<j} \operatorname{Cov}(X_i, X_j).$$

As in our homework, we can combine the definition of correlation (Definition 7.3.4) with the fact that correlation is always in the interval $[-1, 1]$ (Theorem 7.3.5) to obtain

$$-\sigma_i\sigma_j = -\sqrt{V(X_i)V(X_j)} \leq \operatorname{Cov}(X_i, X_j) \leq \sqrt{V(X_i)V(X_j)} = \sigma_i\sigma_j.$$

Therefore $V(T)$ satisfies the bounds

$$\sum_{i=1}^{n} \sigma_i^2 - 2\sum_{i<j} \sigma_i\sigma_j \leq V(T) \leq \sum_{i=1}^{n} \sigma_i^2 + 2\sum_{i<j} \sigma_i\sigma_j.$$

A.F. The assistant appears to be applying the central limit theorem. However, that theorem applies only in the case that the X_i are IID. Regardless of whether the X_i are independent (doubtful, but maybe a useful case to study anyway), they are almost certainly not identically distributed (or else we wouldn't have n means and variances specified). Now, it's possible that T ends up being approximately normal anyway, for reasons unrelated to the central limit theorem. We could assess that possibility by examining old data. In any event, a lot more work is needed, before we can possibly take advantage of the assistant's idea.

B. We follow our usual strategy:

$$\begin{aligned} F_Y(y) &= P(Y \le y) \\ &= P(\log X \le y) \\ &= P(X \le e^y) \\ &= F_X(e^y) \\ \Rightarrow f_Y(y) &= f_X(e^y) \cdot \frac{d}{dy} e^y \\ &= e^y f_X(e^y). \end{aligned}$$

Without knowing more about the support of X, we can't say much about the support of Y. The support is possibly as large as the entire real number line $(-\infty, \infty)$. For example, $Y < 0$ when $X < 1$.

C. If $Y = X$, then $m_{X+Y}(t) = m_{2X}(t) = m_X(2t)$. [Bonus question: Under what conditions on X does this equal $(m_X(t))^2$?]

D.A. As M decreases, it means that we are including more and more earthquakes in our count X, so $E(X)$ should increase. But $X \sim \text{Pois}(\lambda)$, so $E(X) = \lambda$, so λ should increase. [Similarly, if $Y \sim \text{Expo}(\lambda)$ is the inter-arrival time, then $E(Y) = 1/\lambda$ should decrease, again implying that λ should increase.]

D.B. When M is small, λ is large. For the sake of simplicity, round λ to the nearest integer. Then $X \sim \text{Pois}(\lambda)$ can be regarded as $X = X_1 + \cdots X_\lambda$, where the $X_i \sim \text{Pois}(1)$ are IID. By the central limit theorem, $X \approx \text{Norm}(E(X), V(X)) = \text{Norm}(\lambda, \lambda)$. So the PMF of X should be bell-shaped and centered at $x = \lambda$, with inflection points at $x = \lambda \pm \sqrt{\lambda}$. Because X is approximately normal, its skewness and kurtosis should be about 0. [In fact, the skewness is $\lambda^{-1/2}$ and the kurtosis is λ^{-1}.]