STA260 Tutorial 1 Question 2

Question 2

Let $Y_1, Y_2, ..., Y_n$ be n independent observations, but they're **not necessarily identically distributed**. This means it's possible some are from different distributions (normal, exponential, etc.) However, they do conveniently all have the same mean μ and finite variance σ^2. The sample variance is defined as:

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \bar{Y})^2$$

(a) Prove $\mathbb{E}[S^2] = \sigma^2$

~~(b)~~ Prove $\mathbb{E}[S] \leq \sigma$ → requires Jensen's inequality.

(b) Suppose we learned that $Y_1, Y_2, ..., Y_n$ is identically distributed, and they are in fact observations from a **normal** distribution. How convenient! Find the constant a such that: $\mathbb{P}\left(\frac{3S^2}{\sigma^2} \geq a\right) = 0.9$

a) $\mathbb{E}[S^2] = \mathbb{E}\left[\frac{1}{n-1}\sum_{i=1}^{n}(Y_i-\bar{Y})^2\right]$

note that $\sum_{i=1}^{n}(Y_i-\bar{Y})^2 = \sum_{i=1}^{n}(Y_i^2 - 2Y_i\bar{Y} + \bar{Y}^2)$

$= \sum_{i=1}^{n} Y_i^2 - 2\bar{Y}\sum_{i=1}^{n} Y_i + \sum_{i=1}^{n}\bar{Y}^2$

$= \sum_{i=1}^{n} Y_i^2 - 2\bar{Y}(n\bar{Y}) + n\bar{Y}^2$

$= \sum_{i=1}^{n} Y_i^2 - n\bar{Y}^2$

$= \mathbb{E}\left[\frac{1}{n-1}\sum_{i=1}^{n} Y_i^2\right] - \mathbb{E}\left[\frac{1}{n-1} n\bar{Y}^2\right]$

$= \frac{1}{n-1}\sum_{i=1}^{n}\mathbb{E}[Y_i^2] - \frac{n}{n-1}\mathbb{E}[\bar{Y}^2]$

note: $Var(Y_i) = \mathbb{E}(Y_i^2) - \mathbb{E}(Y_i)^2 \Rightarrow \mathbb{E}(Y_i^2) = Var(Y_i) + \mathbb{E}(Y_i)^2$

$= \sigma^2 + \mu^2$

and $\mathbb{E}(\bar{Y}^2) = Var(\bar{Y}) + \mathbb{E}(\bar{Y})^2$

$= \sigma^2/n + \mu^2$

$= \frac{1}{n-1}\sum_{i=1}^{n}(\sigma^2+\mu^2) - \frac{n}{n-1}\left(\frac{\sigma^2}{n} + \mu^2\right)$

$= \frac{n}{n-1}\left[\sigma^2 + \mu^2 - \frac{\sigma^2}{n} - \mu^2\right]$

$= \frac{n}{n-1}\left[\sigma^2\left(\frac{n-1}{n}\right)\right] = \sigma^2$ as desired.

b) Recall: $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$ thus

$n-1=3$ thus from the table: $\chi^2_{0.9,df=3}$: 0.584375

Thus $a = 0.584375$.