Questions

1) (Chernoff Bound)
a) Prove Chernoff bound:

$$\mathbb{P}\{X \geq x\} \leq \min_{t > 0} \frac{\mathbb{E}[e^{tX}]}{e^{tx}}$$ (1)

b) Apply Chernoff bound to Poisson random variable $X$, and show:

$$\mathbb{P}\{X \geq x\} \leq \frac{e^{-\mu}(e\mu)^x}{x^x}, \quad \text{if } x > \mu$$ (2)

$$\mathbb{P}\{X \leq x\} \leq \frac{e^{-\mu}(e\mu)^x}{x^x}, \quad \text{if } x < \mu$$ (3)

where $\mu$ is the mean of $X$.

2) (Balls and Bins) There are $n$ balls thrown independently and uniformly at random into $n$ bins.
a) Show the probability that a particular bin receiving at least $M$ balls is at most $\left(\frac{n}{M}\right)\left(\frac{1}{n}\right)^M$.
b) Show the probability that any bin receiving more than $\frac{3\ln n}{\ln \ln n}$ balls is at most $\frac{1}{n}$.
c) Derive a sufficient number of $n$ that can guarantee that the probability any bin receiving more than 1% of balls is at most 1%.

3) (Bit Strings) An ideal hash function will map a string to an bit-string, such that each bit has a equal probability of being 0 or 1. Suppose that there are a set of strings $S = \{s_1, ..., s_m\}$. We map each string to a fingerprint bit-string using $b$ bits by an ideal hashing function. Then we store the set of fingerprints $F = \{f_1, ..., f_m\}$, which are used to validate whether a new string is a member of $S$ or not.
a) Show that $b = \Omega(\log m)$ bits is necessary for the probability of a false positive being lesser than 1.
b) Show that $b = O(\log m)$ bits is sufficient for the probability of a false positive being at most $\frac{1}{m}$.