Practice questions

Clearly describe the sample space, the events of interest, and the probability model whenever appropriate. In some of the questions you'll need to estimate the quantity of your interest using the method of your choice: Markov's inequality, Chebyshev's inequality, or the Central Limit Theorem. Justify why the method is applicable and discuss the quality of the estimate.

1. Each of the 150 ENGG2430 students shows up to class independently with probability 0.9 and asks Poisson(0.05) questions in there. Let S be the number of students in class and Q the total number of questions asked. Find (a) $\mathrm{E}[S]$, (b) $\mathrm{E}[Q|S]$, (c) $\mathrm{E}[Q]$, (d) $\mathrm{Var}[\mathrm{E}[Q|S]]$, (e) $\mathrm{Var}[Q|S]$, (f) $\mathrm{E}[\mathrm{Var}[Q|S]]$, (g) $\mathrm{Var}[Q]$.

2. 100 people put their hats in a box and each one pulls a random hat out.
   (a) Let G be any 10-person group. What is the probability that everyone in G pulls their own hat?
   (b) What is the expected *number* of 10-person groups in which everyone pulls their own hat?
   (c) Show that the probability that 10 or more people pull their own hat is less than 10^{-6}.

3. In a school fair, you put up a game stall. In each game, the participant pays you \$10, he or she then draws a ball from a box of 9 white balls and 1 red ball, if the ball drawn is red, you pay \$40 back, otherwise the participant gains nothing. Estimate the probability that you have gained \$300 after 100 games.

4. 100 balls are tossed at random into 100 bins. Each ball is equally likely to land in any of the bins, independently of the other balls.
   (a) Find the expected number and variance of the number of non-empty bins.
   (b) Show that there are fewer than 80 non-empty bins with a probability at least 90%.

5. Consider the following simplified model of infection spread. On any given day, any carrier independently infects one additional person with probability p and is cured with probability $1-p$. The number X_d of virus carriers on day d is given by $X_d = 2 \cdot \mathrm{Binomial}(X_{d-1}, p)$.
   (a) Let $e_d = \mathrm{E}[X_d]$. Express e_d in terms of e_{d-1}. What is e_d in terms of X_0, p, and d?
   (b) Show that when $X_0 = 100$ and $p = 0.4$, the probability 100 or more people are carriers on day 21 is less than 1%.
   (c) Let $v_d = \mathrm{Var}[X_d]$. Express v_d in terms of v_{d-1}.
   (d) **(Optional)** Show that when $X_0 = 100$ and $p = 0.6$, the probability that 100 or more people are carriers on day 21 is more than 95%.