AML5153 | Applied Probability and Statistics | Lab Final Exam

1. A machine produces items in batches. For each batch, the machine can be in control or out of control. Suppose the machine is in control for 98% of the production batches; The production defect rate is 0.05 when the machine is in control and 0.2 when the machine is out of control.

   Suppose we want to update our prior knowledge of the probability that the machine is in-/out-of control. The updated probabilites are called posterior probabilities. To that end, we select a small random sample of 5 items from a production batch and inspect how many among them are defective, which is represented by the random variable X. Compute the missing entries in the following table:

Control Status	**Prior Rate**	**Posterior Rate Given $X = j$**					
		$j = 0$	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$
In control	?	?	?	?	?	?	?
Out of control	?	?	?	?	?	?	?

   Suppose we want to interrupt the production process when we are suspicious that the production process is out of control based on the number of defective items we see in a sample of 5. Let us quantify our suspicion as a 50% or greater chance. Would you stop production if 3 defective items are observed?

   Hint: Apply Bayes' theorem to $P(\text{in control} \,|\, X = j)$ and identify what kind of a random variables are $P(X = j \,|\, \text{in control})$ and $P(X = j \,|\, \text{out of control})$.

2. Suppose a random number of K customers shop at a supermarket in a day. Let $X_1, X_2, \ldots, X_K$ represent the random number of items purchased independently by the 1st, 2nd, ..., Kth customer. The total number of items sold in a day is a random number Y such that:

$$Y = X_1 + X_2 + \cdots + X_K.$$

Suppose that on an average 30 customers arrive per day. Each individual customer is

- 20% likely to be in the age group 20-40 (encoded as 0) who buys on an average 10 items based on a Poisson distribution.
- 35% likely to be in the age group 40-60 (encoded as 1) who buys on an average 20 items based on a Poisson distribution.
- 45% likely to be in the age group 60 and above (encoded as 2) who buys on an average 5 items based on a Poisson distribution.

If you are the supermarket owner and want to increase the *expected* total number of items sold in a day. You have two options:

- increase the *expected* number of customers across all age groups by 10% by spending on external advertisement;
- increase the *expected* number of items purchased by customers in the 20-60 age group by 15% by spending on internal (in-shop) advertisement.

(a) Do both options result in an increase in the expected total number of items sold in a day?
Useful functions: `rpois()`, `sample()`, `sum()`, `rep()`, `mean()`.

(b) Which option results in the greatest increase in the expected total number of items sold in a day?

(c) Which option is least risky?
Hint: standard deviation