Use the accompanying spreadsheet as template for completing each of the tasks.

Task A

Retrieve your answers to Question 1 from Homework 5, as well as the solution notes posted on the course website.

What probability did you get for actually having the cancer, given a positive test? How does this compare to the population baseline probability?

The homework sheet follows up with three informal questions: discuss what you would answer to those questions.

Task B

(This is Question 117 from the textbook [Carlton & Devore 2017].)

If the side of a square X is random with the pdf $f_X(x) = x/8$, $0 < x < 4$, and Y is the area of the square, find the pdf of Y.

Task C

A farm grows apples for sale to distributors. The price of apples depends on their weight and distributors will only take apples of a certain weight range. This year's apples from the farm have weights normally distributed with mean 170g and standard deviation 35g. They are sold to distributors in batches of 500.

Distributor A is looking for apples between 120g and 240g in weight: they check for this by taking a random sample of 10 apples from a batch and will accept the batch if at least 8 are in the target weight range.

Distributor B has a different policy. They are interested in any apples weighing over 100g: to test for this they sample 5 and reject the whole batch if even one of those is underweight.

Which distributor is more likely to accept a batch of apples from this farm?

Task D

The Normal distribution can be used as an approximation for calculating probabilities from the Binomial distribution. This task explores how well approximations like this work in a simple case.

If $X \sim \text{Binom}(n, p)$ is a discrete random variable counting successes in n trials each with probability p then we have the following approximation.

$$P(X \leq x) \;=\; B(x; n, p) \;\approx\; \Phi\left(\frac{x + 0.5 - np}{\sqrt{npq}}\right) \qquad x \in \{0, 1, \ldots, n\}$$

This is considered a good approximation in practice if $np \geq 10$ and $nq \geq 10$ where $q = (1 - p)$. Here you will look at a situation where both values are much less than 10.

(a) Suppose discrete random variable $X \sim \text{Binom}(4, 0.5)$ counts the number of heads thrown in four tosses of a fair coin: $X \in \{0, 1, 2, 3, 4\}$. What are the mean and standard deviation of the random variable X?

(b) Use the Normal approximation above to estimate the probability that X takes the value 2.

(c) Complete this to draw up a table of estimates for the probability that X takes each of the possible values 0 to 4.

(d) Now compute a table of the PMF for X showing the exact probability that it takes each of these five possible values.

(e) Calculate the table of errors: what percentage variation is there in each estimate compared to the exact value? Which values have the largest error?