Mathematical Biology Worksheet 2 Solutions

Georgeos Hardo

October 2020

Question 1

Let the RV X be the number of years a part lasts, and let the lifetime of a part exponentially distributed with a mean lifetime of 10 years. We know the expected value of an exponential distribution is $\frac{1}{\lambda}$ (formula book). Therefore 10 years $= \frac{1}{\lambda}$, and so we can write $\lambda = 0.1$ years^{-1}

$$P(X > 7 \text{ years} = \int_7^\infty 0.1e^{-0.1x}dx = 0.5$$

If the expected lifetime of a single part is $\mathbb{E}[X] = 10$ years, then the lifetime of 5 parts is $\mathbb{E}[5X]$. By the linearity of expectation, we can write:

$$\mathbb{E}[5X] = 5\,\mathbb{E}[X] = 50 \text{ years}$$

Question 2

We are given that incubator temperatures are normally distributed with $\mu = 21°$C and $\sigma = 5°$C.

We want to work out the expected production cost of the drug given some conditions.

Let's find the Z-score of 25°C given this non-standard normal distribution:

$$Z = \frac{25 - 21}{5} = 0.8$$

Therefore, the probability (or fraction of time) for which the incubator is $< 25°$C is:

$$\text{Norm}(X < 25; 21, 5) = \text{Norm}(Z < 0.8; 0, 1) = 1 - 0.2119 \text{ (By symmetry - try drawing a diagram)}$$

And the probability (or fraction of time) for which the incubator is $\geq 25°$C is:

$$\text{Norm}(X \geq 25; 21, 5) = \text{Norm}(Z \geq 0.8; 0, 1) = 0.2119$$

And so:

$$\mathbb{E}[\text{price}] = 0.2119 \cdot 30 + (1 - 0.2119) \cdot 25 = 22.19 \text{ GBP}$$

Question 3

We're told a box contains 3 green and 5 red balls, and that we replace them when we choose them.

(a) This can be solved in two ways. There are two possibilities for a sequence of 3 balls which ends with a red ball for the second time: GRR and RGR. And so calculating the probability of getting one of these sequences is trivial:

$$\frac{3}{8} \cdot \frac{5}{8} \cdot \frac{3}{8} + \frac{5}{8} \cdot \frac{3}{8} \cdot \frac{3}{8} = 0.2933$$

Alternatively we can calculate this with a binomial distribution. Let's define picking a red ball as a success. The probability of this event is $\frac{5}{8}$. We can therefore calculate the same probability as follows:

$$P([1 \text{ success in the first two trials}] \cap [\text{Success in final trial}]$$

And because each event is independent:

$$P([1 \text{ success in the first two trials}]) \cdot P([\text{Success in final trial}])$$

The first term can be calculated with a binomial distribution, and the final term is simply $\frac{5}{8}$:

$$\text{Binom}(k = 1; n = 2, p = 5/8) \cdot \frac{5}{8} = 0.2933$$

(b) If we again define picking a red ball as a success, and we want to wait exactly 4 trials before we pick the red ball, then simply:

$$\text{Geom}(k = 4; p = 5/8) = 0.032$$

(c) Given that the probability of picking a red ball is given by a geometric distribution: Geom$(k; p = 5/8)$, and we know the expected value of a geometric distribution is $\frac{1}{p}$ (check your formula book. Then the expected time to wait for a single red ball is $\frac{1}{5/8} = 1.6$ trials. And so the waiting time to get two red balls is twice this: 3.2 trials.

Question 4

This question is quite hard!

(a) The probability that a birthday occurs on any given day of the year i is $P(X_i) = \frac{1}{365}$. Therefore if we perform 100 trials of asking someone: "Is your birthday on day i?" They will answer "yes" with a probability of 1/365. So what is the probability of getting 3 successes (three people having a birthday on day i)?:

$$\text{Binom}(k = 3; n = 100, p = 1/365) = 0.00255$$

There are 365 days in a standard year, and so i runs from 0 to 365. The probability of a birthday is assumed to be the same on every day, and so the expected number of days in the year for which 3 of 100 people will share a birthday is:

$$365 \cdot 0.00255 = 0.93 \text{ days}$$

(b) The expected number of distinct birthdays: let $X_i = 1$ if someone has a birthday on day i, and $X_i = 0$ if they do not. Therefore:

$$P(X_i = 1) = P(\text{At least one birthday on day } i) = 1 - P(\text{No birthday on day } i) = 1 - \left(\frac{364}{365}\right)^{100}$$

This is analogous to saying that there are $1 - \left(\frac{364}{365}\right)^{100}$ on each given day, or that the expected number of birthdays on each day is $1 - \left(\frac{364}{365}\right)^{100}$.

Because $X_i = X_1 = \ldots X_{365}$, the expected number of birthdays in an entire year can be calculated through linearity of expectation:

$$\mathbb{E}[X] = \mathbb{E}[365X_i] = 365\,\mathbb{E}[X_i] = 87.6$$

Question 5

If you want calculus help, ask during the supervision, or plug these integrals into Mathematica by typing: "== Integral if FUNCTION between LOW and HIGH". Mathematica will give you worked solutions!

Given a continuous random variable X has PDF:

$$f(x) = \begin{cases} \frac{4x(9-x^2)}{81} & 0 \le x \le 3 \\ 0 & \text{else} \end{cases}$$

(a) We want verify that $f(x)$ is a valid PDF. A valid PDF has a total area under the curve of 1 between $-\infty < x < \infty$. You can verify this by integrating over the valid range of the function:

$$\int_0^3 \frac{4x(9-x^2)}{81} dx = 1$$

(b) The mean is found by integrating over the valid range the function $xf(x)$:

$$\int_0^3 x\frac{4x(9-x^2)}{81} dx = 1.6$$

(c) Recall that the mode occurs at the maxima of the function (if it has some!) Therefore the mode is given by x where:

$$\frac{\mathrm{d}f(x)}{\mathrm{d}x} = \frac{\mathrm{d}}{\mathrm{d}x}\frac{4x(9-x^2)}{81} = 0 \implies x = \sqrt{3}$$

(d) Recall the median is defined at the point at which the CDF reaches 0.5. This is equivilent to solving for m where:

$$0.5 = \int_0^m \frac{4x(9-x^2)}{81} dx \implies m = 1.62$$

Question 6

(a) We are told X is a CRV with PDF:

$$f(x) = \begin{cases} 4x^3 & 0 \le x \le 0 \\ 0 & \text{else} \end{cases}$$

We are asked to find $P(X \le 2/3 | X > 1/3)$. How can we apply conditional probability to a continuous function? Let us write this first in terms of Bayes rule to switch conditions:

$$P\left(X \le \frac{2}{3} | X > \frac{1}{3}\right) = \frac{P\left(X > \frac{1}{3} | X \le \frac{2}{3}\right) \cdot P\left(X \le \frac{2}{3}\right)}{P\left(X > \frac{1}{3}\right)}$$

It makes more sense to talk in terms of unions when it comes to CDFs, and so we can rewrite this as:

$$P\Big(X \leq \frac{2}{3} | X > \frac{1}{3}\Big) = \frac{P\Big(X > \frac{1}{3} \cup X \leq \frac{2}{3}\Big)}{P\Big(X > \frac{1}{3}\Big)} = \frac{P\Big(\frac{1}{3} < X \leq \frac{2}{3}\Big)}{P\Big(X > \frac{1}{3}\Big)}$$

Solving is as simple as integrating:

$$P\Big(X \leq \frac{2}{3} | X > \frac{1}{3}\Big) = \frac{\int_{\frac{1}{3}}^{\frac{2}{3}} 4x^3 dx}{\int_{\frac{1}{3}}^{1} 4x^3 dx} = \frac{3}{16}$$

(b) We are told X is a CRV with PDF:

$$f(x) = \begin{cases} x^2\Big(2x + \frac{3}{3}\Big) & 0 \leq x \leq 0 \\ 0 & \text{else} \end{cases}$$

We are also told that $Y = \frac{2}{x} + 3$ and are asked to find its variance. Recall the rules for the use of the variance operator:

$$\text{Var}(Y) = \text{Var}\left(\frac{2}{x} + 3\right) = 4\,\text{Var}\left(\frac{1}{x}\right)$$

Given that $\text{Var}\left(\frac{1}{x}\right) = \mathbb{E}[\frac{1}{x^2}] - \mathbb{E}[\frac{1}{x}]^2$, we can calculate these expected values as:

$$\mathbb{E}\left[\frac{1}{x^2}\right] = \int_0^1 x\left(2x + \frac{3}{3}\right)dx = \frac{5}{2}$$
$$\mathbb{E}\left[\frac{1}{x}\right] = \int_0^1 \left(2x + \frac{3}{3}\right)dx = \frac{17}{12}$$

Therefore:

$$\text{Var}\left(\frac{1}{x}\right) = \frac{5}{2} - \frac{17^2}{12^2}$$

And recall that:

$$\text{Var}(Y) = 4\,\text{Var}\left(\frac{1}{x}\right) = \frac{71}{36}$$

Question 7

This question is weirdly worded, but makes sense when you understand it. Only **one** of the locations has food. If the bear does not find food then it has to walk back to where it started and look for food again. It has no memory of where it already looked, and so could choose any path again. Given that each route R_i has an equal probability of 1/3, we can write the expected time as:

$$\mathbb{E}[T] = \mathbb{E}[T|R_1]P(R_1) + \mathbb{E}[T|R_2]P(R_2) + \mathbb{E}[T|R_3]P(R_3)$$

All probabilities are equal and so we can factor them out:

$$\mathbb{E}[T] = \frac{1}{3} \cdot (\mathbb{E}[T|R_1] + \mathbb{E}[T|R_2] + \mathbb{E}[T|R_3])$$

We are told that **the last road** (R_3) leads to the food source. Therefore the expected time to get some food, given the bear takes R_3 is $\mathbb{E}[T|R_3] = 2$ hours. We still do not know what the expected time is, but we know that if the bear takes path R_1 or R_2, it will have wasted that time and not found food. Therefore the expected time to get food if it takes R_1 is $\mathbb{E}[T|R_1] = \mathbb{E}[T] + 1$ hour, and for R_2: $\mathbb{E}[T|R_2] = \mathbb{E}[T] + 6$ hours. We can substitute these in to the previous equation:

$$\mathbb{E}[T] = \frac{1}{3} \cdot ((\mathbb{E}[T] + 1) + (\mathbb{E}[T] + 6) + 2)$$

If you solve for $\mathbb{E}[T]$ you will find that it equals 9 hours.