Problem Section 7

Mon Nov 20, 2023

Learning Outcomes

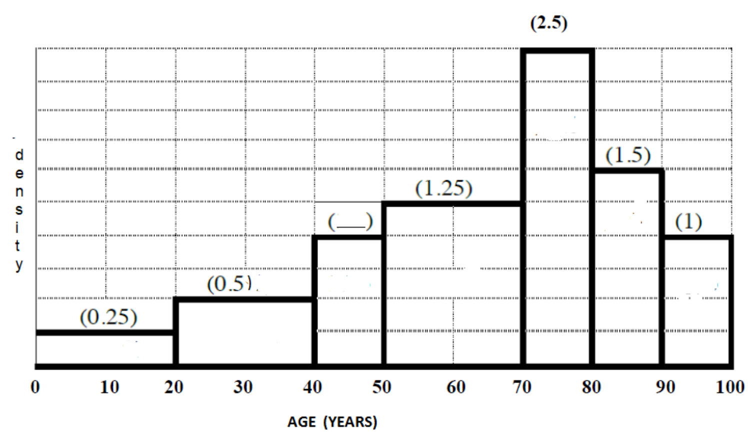
The problems are designed to build conceptual understanding and problem-solving skills. The emphasis is on learning to find, evaluate and build confidence. The specific tasks include:

- Calculate density in a density histogram
- Calculate probabilities in the uniform and exponential distribution
- Apply the law of the unconscious probabilist for a continuous random variable
- Back up and support work with relevant explanations

Exercises

1. The distribution of age at death of a large population is shown as a density histogram below. The density of each bin (reported as percentages) is given in parentheses above the block. For example, the density of the bin 70-80 is 2.5% or 0.0025.

   Find the density of the 40 - 50 bin. Hint: the total area is 1.



The area of the remaining blocks in the histogram is $0.0025 \times 20 + 0.005 \times 20 + 0.0125 \times 20 + 0.025 \times 10 + 0.015 \times 10 + 0.01 \times 10 = 0.9$.

Therefore the area of the block with the missing density must be 0.1. Since the width of the block is 10, that must mean the density (or height) is 0.01 or 1%.

2. Suppose X is uniformly distributed over $[0, 1)$. That is

$$f(x) = 1 \quad 0 \le x < 1.$$

Show that

$$P(-\ln(X) \le 5) = 1 - e^{-5}$$

where ln denotes the natural logarithm function.

$$\begin{aligned}
P(-\ln(X) \le 5) &= P(\ln(X) \ge -5) \\
&= P(e^{\ln(X)} \ge e^{-5}), \\
&= P(X \ge e^{-5}), \\
&= 1 - P(X < e^{-5}), \\
&= 1 - \int_0^{e^{-5}} dx \\
&= 1 - x\Big|_0^{e^{-5}}, \\
&= 1 - e^{-5}.
\end{aligned}$$

3. Among the most famous of all meteor showers are the Perseids, which occur each year in early August. In some areas the frequency of visible Perseids can be as high as forty per hour.
Given that such sightings occur according to a Poisson process, the random variable Y which denotes the time (in hours) between consecutive sightings satisfies $Y \sim Exp(\lambda = 40)$.

a. Calculate the probability that an observer who has just seen a meteor will have to wait at least five minutes before seeing another one. Do the calculation analytically (meaning do the integration) and also in R.

The PDF of Y is given by

$$f(y) = 40e^{-40y} \quad 0 \le y < \infty.$$

We want to find $P(Y > \frac{5}{60})$ that is Y is longer than 5 minutes.

$$\begin{aligned}
P\left(Y > \frac{5}{60}\right) &= \int_{5/60}^{\infty} f(y)dy, \\
&= \int_{5/60}^{\infty} 40e^{-40y}dy, \\
&= -e^{-40y}\big|_{5/60}^{\infty}, \\
&= e^{-5\times40/60}.
\end{aligned}$$

We can calculate this in R as shown below.

```r
pexp(5/60, rate = 40, lower.tail = F)    #P(Y > 5/60)
```

```
## [1] 0.03567399
```

```r
1 - pexp(5/60, rate = 40)                #1 - P(Y <= 5/60)
```

```
## [1] 0.03567399
```

```r
exp(-5*40/60)                          #our calculation
```

```
## [1] 0.03567399
```

Another approach is to define the Poisson random variable X as the number of events in a time interval of 5 minutes. Then $X \sim Pois(40 \times \frac{5}{60})$ since the meteor sightings occur according to a Poisson process. We want to find $P(X = 0)$.

```r
dpois(x = 0, lambda = 5*40/60)
```

```
## [1] 0.03567399
```

b. What is the median length of time (in hours) the observer (from part a) will need to wait before seeing another one? Do the calculation analytically (meaning do the integration) and also in R.

   (Note: In R, the natural logarithm function is `log(x, base = exp(1) )`. The argument `base = exp(1)` is actually the default setting. So you can omit specifying it if you like. For example `log(x)` calculates the natural logarithm of x)

   Here, we want to find the number q so that $P(Y < q) = P(Y > q) = \frac{1}{2}$. Let's find it analytically.

$$\begin{aligned} P(Y < q) &= \int_{-\infty}^{q} f(y)dy \\ &= \int_{0}^{q} 40e^{-40y}dy, \\ &= -e^{-40y}\Big|_0^q \\ &= 1 - e^{-40q} \end{aligned}$$

   We want this probability to equal $\frac{1}{2}$. In other words, we want to solve

$$\begin{aligned} 1 - e^{-40q} &= \frac{1}{2} \\ e^{-40q} &= \frac{1}{2} \\ -40q &= \ln\left(\frac{1}{2}\right) \\ q &= -\frac{1}{40}\ln\left(\frac{1}{2}\right) = 0.017 \end{aligned}$$

   The median waiting time to see the next meteor show is 0.017 hours.

   The calculation is done in R below.

```r
qexp( p = 0.5, rate = 40)
```

```
## [1] 0.01732868
```

4. Suppose X is uniformly distributed on $[a, a+1)$ for some $a > 0$, that is, the PDF of X is:

$$f(x) = 1 \quad a \leq x < a + 1.$$

   Show that

$$E\left(\frac{1}{X}\right) = \ln\left(\frac{(a+1)}{a}\right)$$

   where ln is the natural logarithm function.

By the law of the unconscious probabilist,

$$
\begin{aligned}
E\left[1/X\right] &= \int\limits_{-\infty}^{\infty} \frac{1}{x} f(x) dx, \\
&= \int\limits_{a}^{a+1} \frac{dx}{x} && \text{(by definition of f(x))} \\
&= \ln(x)\big|_{a}^{a+1}, && \text{(ln=natural algorithm)} \\
&= \ln(a+1) - \ln(a), \\
&= \ln\left(\frac{a+1}{a}\right)
\end{aligned}
$$

where the last line follows from the fact that the difference of logarithms is the logarithm of the quotient.