Lab 8 - Continuous Probability Distributions

Carl Gladish

Oct 13, 2025

This lab contains some instructional material along with some questions. You are required to submit your answers as an *html* document produced from an R Notebook. This file will contain *all* your R code *and* your written answers and charts.

Use the R Notebook file Lab_8_Notebook.Rmd as a template to get started. You will need to:

- adjust the author and date fields in the YAML metadata
- complete the missing R chunks below
- type any written answers using R markdown formatting
- "knit" the result to HTML and submit your .html file to Learning Hub

Due date: 11:59pm, two school days from today (weekend days count as half)

Lab Objectives

In previous labs we used the R function `sample` to perform simulations. This approach is suited to simulating outcomes in a *discrete* sample space (i.e., when X is a discrete random variable). In this lab, we will perform simulations when X is a *continuous* random variable.

Once again, we will use the relative frequency approach, which says the probability of an event A is:

$$P(A) \approx \frac{\text{number of times } A \text{ occurs}}{\text{number of trials}}$$

Experiment 1: Waiting for a Bus (Uniform Distribution)

In our first experiment, we imagine a person waiting for a bus that comes reliably every 20 minutes. The person waiting, however, does not know when the previous bus arrived; they only know that the wait time will be between 0 and 20 minutes.

If we define $X =$ the duration of time until the next bus arrives, then X can be modeled as a *continuous* random variable with a *uniform* distribution with minimum $= 0$ and maximum $= 20$.

We can simulate a person waiting for this bus with the `runif` function.

```
X <- runif(n=1, min=0, max=20)
print(X)
```

```
## [1] 7.754091
```

Here, the person waited for 7.7540915 minutes until the next bus arrived.

Note that R rounds the result to an arbitrary number of digits. We can specify the number of significant figures using `options`. Here the output is printed with 4 (or possibly fewer) significant digits.

```
options(digits=4)
runif(n=1, min=0, max=20)
```

```
## [1] 3.254
```

Question 1

For all parts of this question, assume that the wait time X follows a uniform distribution with min 0 and max 20 minutes.

Q1a - Histogram and Probability

Complete each of the two tasks below using both:

- $m = 100$, and
- $m = 10^5$

1. Create a histogram based on m simulated values of X. Your histogram must:
   - use the function `histogram` from the `mosaic` library
   - use arguments `type="density"` and `right=FALSE`
   - have lower class limits $= 0, 1, 2, \ldots$
2. Use the m simulated values of X to determine $P(X < 6)$.

Q1b - Differences

Briefly describe the way(s) in which the two histograms from Q1a differ.

Q1c - Theoretical

Determine the theoretical probability that a person waits less than 6 minutes.

Q1d - Mean and Standard Deviation

Calculate the mean and standard deviation of the wait time X, assuming it is uniformly distributed on the interval $[0, 20]$.

Experiment 2: Still Waiting (Exponential)

Anyone who has waited for a bus knows that no bus comes *exactly* every 20 minutes. More realistically, the time between buses might be anything from 0 to ∞ minutes (with higher numbers being less likely).

An *exponential distribution* with parameter $\beta = 20$ (corresponding to rate $\lambda = \frac{1}{20} = 0.05$) gives a useful model of the time between buses, where the *mean* time is 20 minutes but that time can be anything from 0 to ∞ minutes.

[**Technical point** The exponential distribution is *memory-less*, which means that, in this scenario, it doesn't matter whether the person arrives right after a bus has left or several minutes later. The time X they will wait from that moment has the same probability distribution.]

Question 2

For all parts of this question, assume that the time a person waits for the bus, X, is an *exponential* random variable with $\beta = 20$ minutes.

Q2a - Probabilities

Use `pexp` to find the probability that the time between buses is less than 6 minutes.

Q2b - Probabilities

Calculate each of the following:

- The probability of waiting less than 10 minutes for a bus.
- The probability of waiting more than 15 minutes for a bus.
- The probability of waiting between 5 and 10 minutes for a bus.

Q2c - Percentiles

Use `qexp` to find the median wait time.

Double check that your answer is correct using something like the following:

```
pexp(15, 1/20)
```

```
## [1] 0.5276
```

Q2d - Quartiles

Find the quartiles Q_1 and Q_3 and calculate the interquartile range, IQR for X.

Q2e - Skewness

Is the wait time X significantly skewed? Calculate Pearson's coefficient of skewness to support your answer.

Q2f - Histogram

Use `rexp` to sample m values of X (m given below) and then generate a histogram showing the distribution of wait times, X. Do this for each of the following values of m. Afterwards, make an observation about the difference(s) between the two histograms.

- $m = 100$ trials
- $m = 10^5$ trials

Q2g - Simulating X

Using $m = 10^5$ trials, find the probability that a person waits less than 6 minutes.

[Note: this is the simulation version of Q2a, so your result should be very similar.]

Q2h - Exact Calculation

Use a formula (not R) to find the probability that a person will wait less than 6 minutes for a bus.

Experiment 3: Quality Control (Normal Distribution)

In mass-production, companies aim to produce large quantities of identical goods. In practice, the product units exhibit random variability that is typically modeled as a normal distribution.

For example, a battery manufacturer produces thousands of 9V batteries. Ideally, each of the batteries have a measured voltage of exactly $9.0000000\,V$. In practice, however, there is some variation in the measured voltages. For instance, it may be that the actual measured voltages of batteries manufactured by this company follow a normal distribution with mean $\mu = 9.01\,V$ and standard deviation $\sigma = 0.05\,V$.

We use `pnorm` to calculate cumulative probabilities for a normally distributed variable. For instance, if X is the battery voltage described above, then $P(X < 9.07)$ is given by

```
pnorm(9.07, mean=9.01, sd=0.05)
```

```
## [1] 0.8849
```

Question 3

For all parts of this question, assume that the measured voltage, X, of batteries manufactured by a company follows a normal distribution with mean $\mu = 9.01\,V$ and standard deviation $\sigma = 0.05\,V$.

Q3a - Probabilities

- Find the probability that X is less than $9.03\,V$
- Find the probability that X exceeds $9.02\,V$
- Find the probability that X is between $8.9\,V$ and $9.1\,V$

Percentiles of a Normal Distribution

We use the function `qnorm` to do the "inverse" of `pnorm`. The expression

```
qnorm(p, mean = mu, sd = sigma)
```

gives the value x such that $P(X \leq x) = p$.

For instance, when a random variable follows a normal distribution with mean 9.01 and standard deviation 0.05, there is a 75% chance that a randomly-chosen value will be less than 9.044.

```
qnorm(0.75, mean=9.01, sd=0.05)
```

```
## [1] 9.044
```

Q3b - Battery Voltage Percentiles

- Find the voltage that is larger than 95% of measured voltages.
- Find the voltage that is lower than 95% of measured voltages.
- Find the 25th percentile voltage.

Sampling From a Normal Distribution

We can use the function `rnorm` to generate random samples (or "observations") of X when X follows a normal distribution with a given mean and standard deviation. For instance, to get 100 values of X, use:

```
rnorm(100, mean=mu, sd=sigma)
```

Q3c - Simulation

Suppose the battery manufacturer will ship (i.e., sell) batteries whose measured voltages are between $8.9\,V$ and $9.1\,V$. Give a command that returns the probability that a randomly selected battery can be shipped, based on:

- $m = 100$ simulated trials
- $m = 10^5$ simulated trials

[Note: this a simulation version of Q3a, so your answers should be very similar.]

Q3d - Theory

Now use the Z-table approach (from lecture) to find the proportion of manufactured batteries that can be shipped.