**HW 4 supplement: exponential distribution**

**This supplement outlines how to use the** rexp **function, which will be useful in HW4. You do not have to hand it in.**

In Lab 6, we modelled the waiting times for a bus that comes every 20 minutes. Anyone who has waited for buses knows that no bus comes *exactly* every 20 minutes (or exactly every 5 minutes, or exactly every hour). More realistically, a bus may come *on average* every 20 minutes. The exponential distribution with mean 20 gives a fairly accurate model of waiting times **between** buses.

We use the **qexp()** function to model exponential waiting times. For an exponential distribution with rate **r,**  the command

qexp(p, rate = r, lower.tail = TRUE, log.p = FALSE)

returns the value such that there is a probability p that the time between occurrences will be **less** than that value.

> qexp(.1, rate = 1, lower.tail = TRUE, log.p = FALSE)

[1] 0.10536

> qexp(.5, rate = 1, lower.tail = TRUE, log.p = FALSE)

[1] 0.69315

For instance, at a mean rate of 1 occurrence per minute, there is a 50% chance the waiting time between occurrences will be less than 0.69315 minutes.

1. For the bus example, what is **r**? (Hint: the units of **r** are “buses per minute”.)
2. Find qexp(.5, rate = r, lower.tail = TRUE, log.p = FALSE) for that value of **r**. This should be the median of an exponentially-distributed random variable with mean 20. How does it compare to the mean? Based on the distribution of the exponential random variable, does this makes sense? (Hint: the word “skew” should appear in your answer.)

In order to model **n** waiting times, we can use the **rexp** command to generate random values that follow an exponential distribution. For instance, the command

> rexp(n, r)

gives **n** values that follow an exponential distribution with rate **r**.

1. Generate appropriately-labelled histograms that give the frequency of waits between buses that have exponentially-distributed waiting times with mean 20 minutes for n=100, 1000, and 10000. Use the same number of classes for each. Do the distributions look exponential? Now generate three additional histograms for that give the frequency of waits between buses that have exponentially-distributed waiting times with mean 5 minutes for n=100, 1000, and 10000. How do these histograms compare to the last ones?